MATHEMATICAL BIOLOGY

Jeffrey R. Chasnov
Hong Kong University of Science and Technology

Hong Kong University of Science and Technology

Mathematical Biology (Chasnov)

Jeffrey R. Chasnov

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.

LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

The LibreTexts name, logo, and book covers are trademarked by LibreTexts, Inc. (a 501(c)(3) not-for-profit organization) and protected against unauthorized use.

This text was compiled on 08/01/2026

- **InfoPage** by Delmar Larsen is licensed Public Domain.

TABLE OF CONTENTS

Licensing

Preface

1: Population Dynamics

- 1.1: The Malthusian Growth Model
- 1.2: The Logistic Equation
- 1.3: A Model of Species Competition
- 1.4: The Lotka-Volterra Predator-Prey Model

2: Age-structured Populations

- 2.1: Fibonacci's Rabbits
- 2.2: The Golden Ratio Φ
- 2.3: The Fibonacci Numbers in a Sunflower
- 2.4: Rabbits are an Age-Structured Population
- 2.5: Discrete Age-Structured Populations
- 2.6: Continuous Age-Structured Populations
- 2.7: The Brood Size of a Hermaphroditic Worm

3: Stochastic Population Growth

- 3.1: A Stochastic Model of Population Growth
- 3.2: Asymptotics of Large Initial Populations
- 3.3: Simulation of Population Growth

4: Infectious Disease Modeling

- 4.1: The SI Model
- 4.2: The SIS Model
- 4.3: The SIR Epidemic Disease Model
- 4.4: Vaccination
- 4.5: The SIR Endemic Disease Model
- 4.6: Evolution of Virulence

5: Population Genetics

- 5.1: Haploid Genetics
- 5.2: Diploid Genetics
- 5.3: Frequency-Dependent Selection
- 5.4: Linkage Equilibrium
- 5.5: Random Genetic Drift

6: Biochemical Reactions

- 6.1: The Law of Mass Action
- 6.2: Enzyme Kinetics
- 6.3: Competitive Inhibition
- 6.4: Allosteric Inhibition
- 6.5: Cooperativity

7: Sequence Alignment

- 7.1: DNA
- 7.2: Brute Force Alignment
- 7.3: Dynamic Programming
- 7.4: Gaps
- 7.5: Local Alignments
- 7.6: Software

Index

Glossary

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

Preface

What follows are my lecture notes for Math 4333: Mathematical Biology, taught at the Hong Kong University of Science and Technology. This applied mathematics course is primarily for final year mathematics major and minor students. Other students are also welcome to enroll, but must have the necessary mathematical skills.

My main emphasis is on mathematical modeling, with biology the sole application area. I assume that students have no knowledge of biology, but I hope that they will learn a substantial amount during the course. Students are required to know differential equations and linear algebra, and this usually means having taken two courses in these subjects. I also touch on topics in stochastic modeling, which requires some knowledge of probability. A full course on probability, however, is not a prerequisite though it might be helpful.

Biology, as is usually taught, requires memorizing a wide selection of facts and remembering them for exams, sometimes forgetting them soon after. For students exposed to biology in secondary school, my course may seem like a different subject. The ability to model problems using mathematics requires almost no rote memorization, but it does require a deep understanding of basic principles and a wide range of mathematical techniques. Biology offers a rich variety of topics that are amenable to mathematical modeling, and I have chosen specific topics that I have found to be the most interesting.

If, as a UST student, you have not yet decided if you will take my course, please browse these lecture notes to see if you are interested in these topics. Other web surfers are welcome to download these notes from

\[\text { http://www.math.ust.hk/ machas/mathematical-biology.pdf } \nonumber \]

and to use them freely for teaching and learning. I welcome any comments, suggestions, or corrections sent to me by email (jeffrey.chasnov@ust.hk). Although most of the material in my notes can be found elsewhere, I hope that some of it will be considered to be original.

JEFFREY R. CHASNOV

Hong Kong

May 2009

CHAPTER OVERVIEW

1: Population Dynamics

Populations grow in size when the birth rate exceeds the death rate. Thomas Malthus, in An Essay on the Principle of Population (1798), used unchecked population growth to famously predict a global famine unless governments regulated family size-an idea later echoed by Mainland China's one-child policy. The reading of Malthus is said by Charles Darwin in his autobiography to have inspired his discovery of what is now the cornerstone of modern biology: the principle of evolution by natural selection.

The Malthusian growth model is the granddaddy of all population models, and we begin this chapter with a simple derivation of the famous exponential growth law. Unchecked exponential growth obviously does not occur in nature, and population growth rates may be regulated by limited food or other environmental resources, and by competition among individuals within a species or across species. We will develop models for three types of regulation. The first model is the wellknown logistic equation, a model that will also make an appearance in subsequent chapters. The second model is an extension of the logistic model to species competition. And the third model is the famous Lotka-Volterra predator-prey equations. Because all these mathematical models are nonlinear differential equations, mathematical methods to analyze such equations will be developed.

1.1: The Malthusian Growth Model

1.2: The Logistic Equation

1.3: A Model of Species Competition

1.4: The Lotka-Volterra Predator-Prey Model

This page titled 1: Population Dynamics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

1.1: The Malthusian Growth Model

Let $N(t)$ be the number of individuals in a population at time t, and let b and d be the average per capita birth rate and death rate, respectively. In a short time Δt, the number of births in the population is $b\Delta tN$, and the number of deaths is $d\Delta tN$. An equation for N at time $t+\Delta t$ is then determined to be

$$N(t+\Delta t) = N(t) + b\Delta tN(t) - d\Delta tN(t)$$

which can be rearranged to

$$\frac{N(t+\Delta t) - N(t)}{\Delta t} = (b-d)N(t)$$

and as $\Delta t \to 0$

$$\frac{dN}{dt} = (b-d)N.$$

With an initial population size of N_0, and with $r = b - d$ positive, the solution for $N = N(t)$ grows exponentially:

$$N(t) = N_0 e^{rt}$$

With population size replaced by the amount of money in a bank, the exponential growth law also describes the growth of an account under continuous compounding With interest rate r.

This page titled 1.1: The Malthusian Growth Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

1.2: The Logistic Equation

The Logistic equation

The exponential growth law for population size is unrealistic over long times. Eventually, growth will be checked by the over-consumption of resources. We assume that the environment has an intrinsic carrying capacity K, and populations larger than this size experience heightened death rates.

To model population growth with an environmental carrying capacity K, we look for a nonlinear equation of the form

$$\frac{dN}{dt} = rNF(N)$$

where $F(N)$ provides a model for environmental regulation. This function should satisfy $F(0) = 1$ (the population grows exponentially with growth rate r when N is small), $F(K) = 0$ (the population stops growing at the carrying capacity), and $F(N) < 0$ when $N > K$ (the population decays when it is larger than the carrying capacity). The simplest function $F(N)$ satisfying these conditions is linear and given by $F(N) = 1 - N/K$. The resulting model is the well-known logistic equation,

$$\frac{dN}{dt} = rN(1 - N/K)$$

an important model for many processes besides bounded population growth.

Although (1.2.2) is a nonlinear equation, an analytical solution can be found by separating the variables. Before we embark on this algebra, we first illustrate some basic concepts used in analyzing nonlinear differential equations.

Fixed points, also called equilibria, of a differential equation such as (1.2.2) are defined as the values of N where $dN/dt = 0$. Here, we see that the fixed points of (1.2.2) are $N = 0$ and $N = K$. If the initial value of N is at one of these fixed points, then N will remain fixed there for all time. Fixed points, however, can be stable or unstable. A fixed point is stable if a small perturbation from the fixed point decays to zero so that the solution returns to the fixed point. Likewise, a fixed point is unstable if a small perturbation grows exponentially so that the solution moves away from the fixed point. Calculation of stability by means of small perturbations is called linear stability analysis. For example, consider the general one-dimensional differential equation (using the notation $\dot{x} = dx/dt$)

$$\dot{x} = f(x)$$

with x_* a fixed point of the equation, that is $f(x_*) = 0$. To determine analytically if x_* is a stable or unstable fixed point, we perturb the solution. Let us write our solution $x = x(t)$ in the form

$$x(t) = x_* + \epsilon(t)$$

where initially $\epsilon(0)$ is small but different from zero. Substituting (1.2.4) into (1.2.3), we obtain

$$\begin{aligned} \dot{\epsilon} &= f(x_* + \epsilon) \\ &= f(x_*) + \epsilon f'(x_*) + \dots \\ &= \epsilon f'(x_*) + \dots , \end{aligned}$$

where the second equality uses a Taylor series expansion of $f(x)$ about x_* and the third equality uses $f(x_*) = 0$. If $f'(x_*) \neq 0$, we can neglect higher-order terms in ϵ

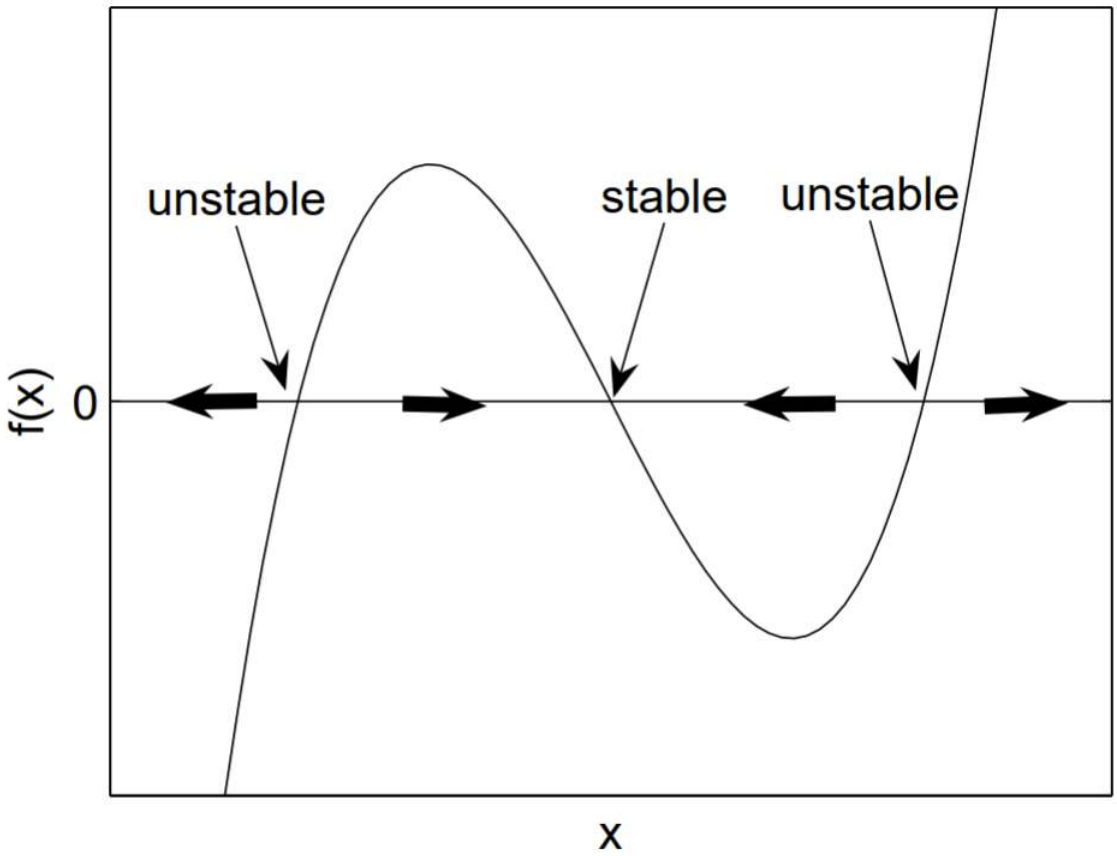


Figure 1.1: Determining one-dimensional stability using a graphical approach.

for small times, and integrating we have

$$\epsilon(t) = \epsilon(0)e^{f'(x_*)t}$$

The perturbation $\epsilon(t)$ to the fixed point x_* goes to zero as $t \to \infty$ provided $f'(x_*) < 0$. Therefore, the stability condition on x_* is

$$x_* \text{ is } \begin{cases} \text{a stable fixed point if} & f'(x_*) < 0, \\ \text{an unstable fixed point if} & f'(x_*) > 0. \end{cases}$$

Another equivalent but sometimes simpler approach to analyzing the stability of the fixed points of a one-dimensional nonlinear equation such as (1.2.3) is to plot $f(x)$ versus x. We show a generic example in Fig. 1.1. The fixed points are the x-intercepts of the graph. Directional arrows on the x-axis can be drawn based on the sign of $f(x)$. If $f(x) < 0$, then the arrow points to the left; if $f(x) > 0$, then the arrow points to the right. The arrows show the direction of motion for a particle at position x satisfying $\dot{x} = f(x)$. As illustrated in Fig. 1.1, fixed points with arrows on both sides pointing in are stable, and fixed points with arrows on both sides pointing out are unstable.

In the logistic equation (1.2.2), the fixed points are $N_* = 0, K$. A sketch of $F(N) = rN(1 - N/K)$ versus N, with $r, K > 0$ in Fig. 1.2 immediately shows that $N_* = 0$ is an unstable fixed point and $N_* = K$ is a stable fixed point. The analytical approach computes $F'(N) = r(1 - 2N/K)$, so that $F'(0) = r > 0$ and $F'(K) = -r < 0$. Again we conclude that $N_* = 0$ is unstable and $N_* = K$ is stable.

We now solve the logistic equation analytically. Although this relatively simple equation can be solved as is, we first nondimensionalize to illustrate this very important technique that will later prove to be most useful. Perhaps here one can guess the appropriate unit of time to be $1/r$ and the appropriate unit of population size to be K. However, we prefer to demonstrate a more general technique that may be usefully applied to equations for which the appropriate dimensionless variables are difficult to guess. We begin by nondimensionalizing time and population size:

$$\tau = t/t_*, \quad \eta = N/N_*$$

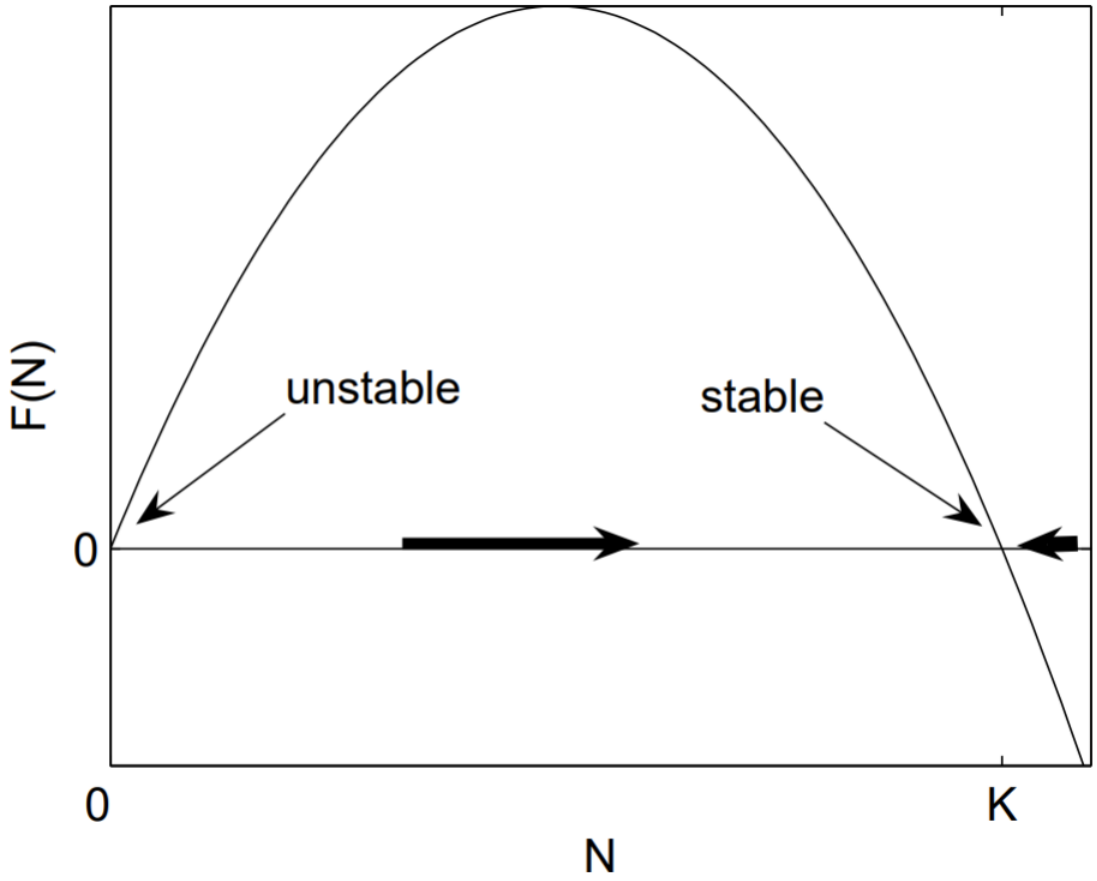


Figure 1.2: Determining stability of the fixed points of the logistic equation.

where t_* and N_* are unknown dimensional units. The derivative N is computed as

$$\frac{dN}{dt} = \frac{d\,(N_*\eta)}{d\tau}\frac{d\tau}{dt} = \frac{N_*}{t_*}\frac{d\eta}{d\tau}$$

Therefore, the logistic equation (1.2.2) becomes

$$\frac{d\eta}{d\tau} = rt_*\eta\left(1 - \frac{N_*\eta}{K}\right),$$

which assumes the simplest form with the choices $t_* = 1/r$ and $N_* = K$. Therefore, our dimensionless variables are

$$\tau = rt, \quad \eta = N/K$$

and the logistic equation, in dimensionless form, becomes

$$\frac{d\eta}{d\tau} = \eta(1-\eta),$$

with the dimensionless initial condition $\eta(0) = \eta_0 = N_0/K$, where N_0 is the initial population size. Note that the dimensionless logistic equation (1.4) has no free parameters, while the dimensional form of the equation (1.2.2) contains r and K. Reduction in the number of free parameters (here, two: r and K) by the number of independent units (here, also two: time and population size) is a general feature of nondimensionalization. The theoretical result is known as the Buckingham Pi Theorem. Reducing the number of free parameters in a problem to the absolute minimum is especially important before proceeding to a numerical solution. The parameter space that must be explored may be substantially reduced.

Solving the dimensionless logistic equation (1.4) can proceed by separating the variables. Separating and integrating from $\tau = 0$ to τ and η_0 to η yields

$$\int_{\eta_0}^{\eta} \frac{d\eta}{\eta(1-\eta)} = \int_0^{\tau} d\tau$$

The integral on the left-hand-side can be performed using the method of partial fractions:

$$\begin{aligned}\frac{1}{\eta(1-\eta)} &= \frac{A}{\eta} + \frac{B}{1-\eta} \\ &= \frac{A + (B-A)\eta}{\eta(1-\eta)}\end{aligned}$$

and by equating the coefficients of the numerators proportional to η^0 and η^1, we find that $A = 1$ and $B = 1$. Therefore,

$$\begin{aligned}\int_{\eta_0}^{\eta} \frac{d\eta}{\eta(1-\eta)} &= \int_{\eta_0}^{\eta} \frac{d\eta}{\eta} + \int_{\eta_0}^{\eta} \frac{d\eta}{(1-\eta)} \\ &= \ln\frac{\eta}{\eta_0} - \ln\frac{1-\eta}{1-\eta_0} \\ &= \ln\frac{\eta\,(1-\eta_0)}{\eta_0(1-\eta)} \\ &= \tau\end{aligned}$$

Solving for η, we first exponentiate both sides and then isolate η :

$$\frac{\eta\,(1-\eta_0)}{\eta_0(1-\eta)} = e^{\tau}, \text{ or } \quad \eta\,(1-\eta_0) = \eta_0 e^{\tau} - \eta\eta_0 e^{\tau}$$

$$\text{or } \eta\,(1-\eta_0+\eta_0 e^{\tau}) = \eta_0 e^{\tau}, \text{ or } \quad \eta = \frac{\eta_0}{\eta_0 + (1-\eta_0)\,e^{-\tau}}$$

Returning to the dimensional variables, we finally have

$$N(t) = \frac{N_0}{N_0/K + (1 - N_0/K)\,e^{-rt}}$$

There are several ways to write the final result given by (1.5). The presentation of a mathematical result requires a good aesthetic sense and is an important element of mathematical technique. When deciding how to write (1.5), I considered if it was easy to observe the following limiting results: (1) $N(0) = N_0$; (2) $\lim_{t\to\infty} N(t) = K$; and (3) $\lim_{K\to\infty} N(t) = N_0 \exp(rt)$

In Fig. 1.3, we plot the solution to the dimensionless logistic equation for initial conditions $\eta_0 = 0.02, 0.2, 0.5, 0.8, 1.0$, and 1.2. The lowest curve is the characteristic 'S-shape' usually associated with the solution of the logistic equation. This sigmoidal curve appears in many other types of models. The MATLAB script to produce Fig. 1.3 is shown below.

```
eta0=[0.02 .2 .5 .8 1 1.2];
tau=linspace(0,8);
for i=1:length(eta0)
    eta=eta0(i)./(eta0(i)+(1-eta0(i)).*exp(-tau));
    plot(tau,eta);hold on
end
axis([0 8 0 1.25]);
xlabel('\tau'); ylabel('\eta'); title('Logistic Equation');
```

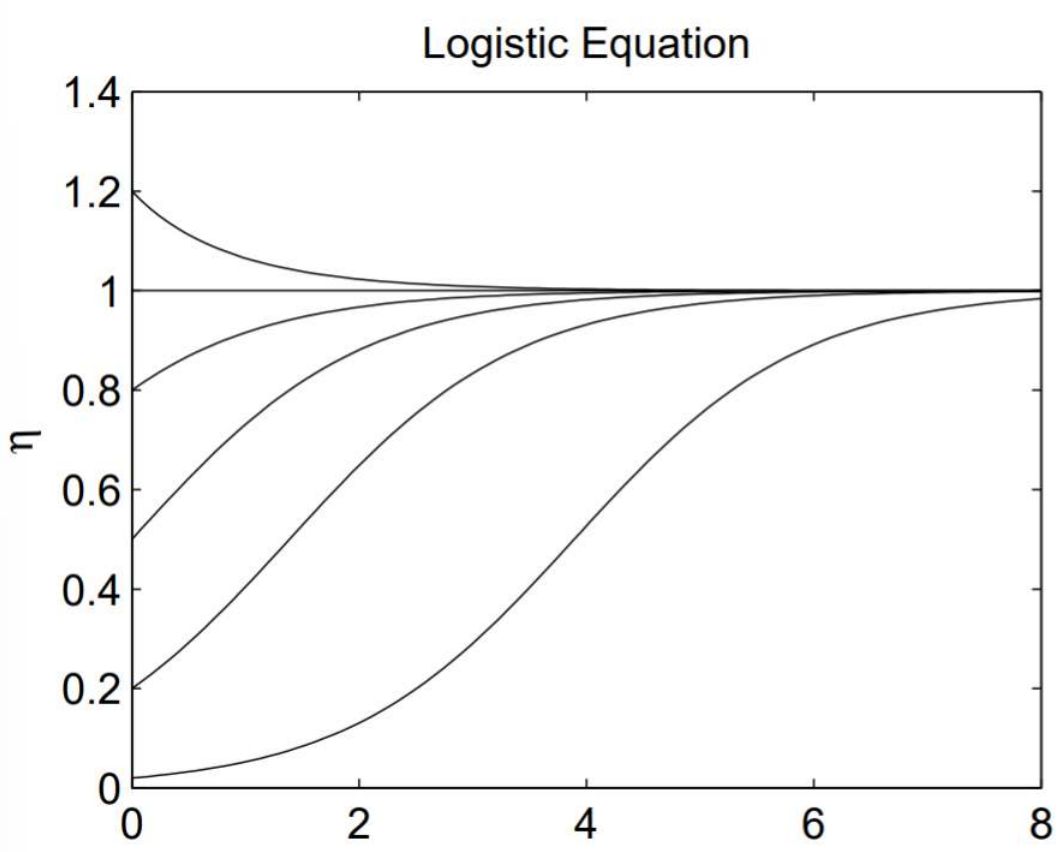


Figure 1.3: Solutions of the dimensionless logistic equation

This page titled 1.2: The Logistic Equation is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

1.3: A Model of Species Competition

A model of species competition

Suppose that two species compete for the same resources. To build a model, we can start with logistic equations for both species. Different species would have different growth rates and different carrying capacities. If we let N_1 and N_2 be the number of individuals of species one and species two, then

$$\frac{dN_1}{dt} = r_1 N_1 \left(1 - N_1/K_1\right)$$
$$\frac{dN_2}{dt} = r_2 N_2 \left(1 - N_2/K_2\right)$$

These are uncoupled equations so that asymptotically, $N_1 \to K_1$ and $N_2 \to K_2$. How do we model the competition between species? If N_1 is much smaller than K_1, and N_2 much smaller than K_2, then resources are plentiful and populations grow exponentially with growth rates r_1 and r_2. If species one and two compete, then the growth of species one reduces resources available to species two, and viceversa. Since we do not know the impact species one and two have on each other, we introduce two additional parameters to model the competition. A reasonable modification that couples the two logistic equations is

$$\frac{dN_1}{dt} = r_1 N_1 \left(1 - \frac{N_1 + \alpha_{12} N_2}{K_1}\right), \tag{1.3.1}$$
$$\frac{dN_2}{dt} = r_2 N_2 \left(1 - \frac{\alpha_{21} N_1 + N_2}{K_2}\right), \tag{1.3.2}$$

where α_{12} and α_{21} are dimensionless parameters that model the consumption of species one's resources by species two, and vice-versa. For example, suppose that both species eat exactly the same food, but species two consumes twice as much as species one. Since one individual of species two consumes the equivalent of two individuals of species one, the correct model is $\alpha_{12} = 2$ and $\alpha_{21} = 1/2$. Another example supposes that species one and two occupy the same niche, consume resources at the same rate, but may have different growth rates and carrying capacities. Can the species coexist, or does one species eventually drive the other to extinction? It is possible to answer this question without actually solving the differential equations. With $\alpha_{12} = \alpha_{21} = 1$ as appropriate for this example, the coupled logistic equations (1.3.1 and 1.3.2) become

$$\frac{dN_1}{dt} = r_1 N_1 \left(1 - \frac{N_1 + N_2}{K_1}\right), \quad \frac{dN_2}{dt} = r_2 N_2 \left(1 - \frac{N_1 + N_2}{K_2}\right)$$

For sake of argument, we assume that $K_1 > K_2$. The only fixed points other than the trivial one $(N_1, N_2) = (0, 0)$ are $(N_1, N_2) = (K_1, 0)$ and $(N_1, N_2) = (0, K_2)$. Stability can be computed analytically by a two-dimensional Taylor-series expansion, but here a simpler argument can suffice. We first consider $(N_1, N_2) = (K_1, \epsilon)$, with ϵ small. Since $K_1 > K_2$, observe from (1.3.3) that $\dot{N}_2 < 0$ so that species two goes extinct. Therefore $(N_1, N_2) = (K_1, 0)$ is a stable fixed point. Now consider $(N_1, N_2) = (\epsilon, K_2)$, with ϵ small. Again, since $K_1 > K_2$, observe from (1.3.3) that $\dot{N}_1 > 0$ and species one increases in number. Therefore, $(N_1, N_2) = (0, K_2)$ is an unstable fixed point. We have thus found that, within our coupled logistic model, species that occupy the same niche and consume resources at the same rate cannot coexist and that the species with the largest carrying capacity will survive and drive the other species to extinction. This is the so-called principle of competitive exclusion, also called K-selection since the species with the largest carrying capacity wins. In fact, ecologists also talk about r-selection; that is, the species with the largest growth rate wins. Our coupled logistic model does not model r-selection, demonstrating the potential limitations of a too simple mathematical model.

For some values of α_{12} and α_{21}, our model admits a stable equilibrium solution where two species coexist. The calculation of the fixed points and their stability is more complicated than the calculation just done, and I present only the results. The stable coexistence of two species within our model is possible only if $\alpha_{12} K_2 < K_1$ and $\alpha_{21} K_1 < K_2$.

This page titled 1.3: A Model of Species Competition is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

1.4: The Lotka-Volterra Predator-Prey Model

1.4 The Lotka-Volterra predator-prey model

Pelt-trading records (Fig. 1.4) of the Hudson Bay company from over almost a century display a near-periodic oscillation in the number of trapped snowshoe hares and lynxes. With the reasonable assumption that the recorded number of trapped animals is proportional to the animal population, these records suggest that predator-prey populations-as typified by the hare and the lynx-can oscillate over time. Lotka and Volterra independently proposed in the 1920 s a mathematical model for the population dynamics of a predator and prey, and these Lotka-Volterra predator-prey equations have since become an iconic model of mathematical biology.

To develop these equations, suppose that a predator population feeds on a prey population. We assume that the number of prey grow exponentially in the absence of predators (there is unlimited food available to the prey), and that the number of predators decay exponentially in the absence of prey (predators must eat prey or starve). Contact between predators and prey increases the number of predators and decreases the number of prey.

Let $U(t)$ and $V(t)$ be the number of prey and predators at time t. To develop a coupled differential equation model, we consider population sizes at time $t + \Delta t$.

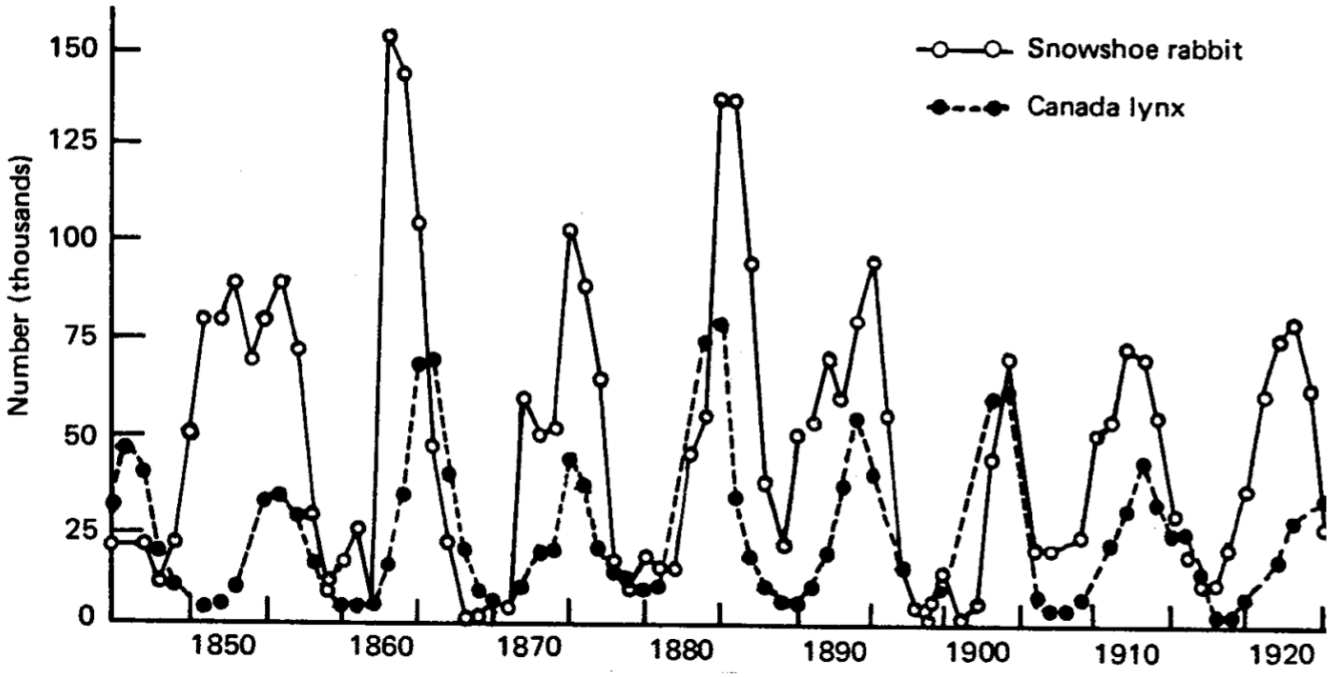


Figure 1.4: Pelt-trading records of the Hudson Bay Company for the snowshoe hare and its predator the lynx. [From E.P. Odum, Fundamentals of Ecology, 1953.]

Exponential growth of prey in the absence of predators and exponential decay of predators in the absence of prey can be modeled by the usual linear terms. The coupling between prey and predator must be modeled with two additional parameters. We write the population sizes at time $t + \Delta t$ as

$$U(t + \Delta t) = U(t) + \alpha \Delta t U(t) - \gamma \Delta t U(t) V(t)$$
$$V(t + \Delta t) = V(t) + e\gamma \Delta t U(t) V(t) - \beta \Delta t V(t)$$

The parameters α and β are the average per capita birthrate of the prey and the deathrate of the predators, in the absence of the other species. The coupling terms model contact between predators and prey. The parameter γ is the fraction of prey caught per predator per unit time; the total number of prey caught by predators during time Δt is $\gamma \Delta t UV$. The prey eaten is then converted into newborn predators (view this as a conversion of biomass), with conversion factor e, so that the number of predators during time Δt increases by e\gamma $\Delta t UV$.

Converting these equations into differential equations by letting $\Delta t \to 0$, we obtain the well-known Lotka-Volterra predator-prey equations

$$\frac{dU}{dt} = \alpha U - \gamma UV, \quad \frac{dV}{dt} = e\gamma UV - \beta V$$

Before analyzing the Lotka-Volterra equations, we first review fixed point and linear stability analysis applied to what is called an autonomous system of differential equations. For simplicity, we consider a system of only two differential equations of the form

$$\dot{x} = f(x, y), \quad \dot{y} = g(x, y),$$

though our results can be generalized to larger systems. The system given by (1.4.2) is said to be autonomous since f and g do not depend explicitly on the independent variable t. Fixed points of this system are determined by setting $\dot{x} = \dot{y} = 0$ and solving for x and y. Suppose that one fixed point is (x_*, y_*). To determine its linear stability, we consider initial conditions for (x, y) near the fixed point with small independent perturbations in both directions, i.e., $x(0) = x_* + \epsilon(0), y(0) = y_* + \delta(0)$. If the initial perturbation grows in time, we say that the fixed point is unstable; if it decays, we say that the fixed point is stable. Accordingly, we let

$$x(t) = x_* + \epsilon(t), \quad y(t) = y_* + \delta(t),$$

and substitute $(1.4.3)$ into $(1.4.2)$ to determine the time-dependence of ϵ and δ. Since x_* and y_* are constants, we have

$$\dot{\epsilon} = f\left(x_* + \epsilon, y_* + \delta\right), \quad \dot{\delta} = g\left(x_* + \epsilon, y_* + \delta\right)$$

The linear stability analysis proceeds by assuming that the initial perturbations $\epsilon(0)$ and $\delta(0)$ are small enough to truncate the two-dimensional Taylor-series expansion of f and g about $\epsilon = \delta = 0$ to first-order in ϵ and δ. Note that in general, the two-dimensional Taylor series of a function $F(x, y)$ about the origin is given by

$$F(x, y) = F(0, 0) + xF_x(0, 0) + yF_y(0, 0) + \frac{1}{2}\left[x^2 F_{xx}(0, 0) + 2xyF_{xy}(0, 0) + y^2 F_{yy}(0, 0)\right] + ...$$

where the terms in the expansion can be remembered by requiring that all of the partial derivatives of the series agree with that of $F(x, y)$ at the origin. We now Taylor-series expand $f\left(x_* + \epsilon, y_* + \delta\right)$ and $g\left(x_* + \epsilon, y_* + \delta\right)$ about $(\epsilon, \delta) = (0, 0)$. The constant terms vanish since (x_*, y_*) is a fixed point, and we neglect all terms with higher orders than ϵ and δ. Therefore,

$$\dot{\epsilon} = \epsilon f_x\left(x_*, y_*\right) + \delta f_y\left(x_*, y_*\right), \quad \dot{\delta} = \epsilon g_x\left(x_*, y_*\right) + \delta g_y\left(x_*, y_*\right),$$

which may be written in matrix form as

$$\frac{d}{dt}\begin{pmatrix} \epsilon \\ \delta \end{pmatrix} = \begin{pmatrix} f_x^* & f_y^* \\ g_x^* & g_y^* \end{pmatrix}\begin{pmatrix} \epsilon \\ \delta \end{pmatrix}$$

where $f_x^* = f_x\left(x_*, y_*\right)$, etc. Equation (1.4.6) is a system of linear ode's, and its solution proceeds by assuming the form

$$\begin{pmatrix} \epsilon \\ \delta \end{pmatrix} = e^{\lambda t}\mathbf{v}$$

Upon substitution of $(1.4.7)$ into $(1.4.6)$, and canceling $e^{\lambda t}$, we obtain the linear algebra eigenvalue problem

$$\mathrm{J}^*\mathbf{v} = \lambda\mathbf{v}, \text{ with } \mathrm{J}^* = \begin{pmatrix} f_x^* & f_y^* \\ g_x^* & g_y^* \end{pmatrix}$$

where λ is the eigenvalue, $\mathbf{v}$ the corresponding eigenvector, and J^* the Jacobian matrix evaluated at the fixed point. The eigenvalue is determined from the characteristic equation

$$\det\left(J^* - \lambda I\right) = 0,$$

which for a two-by-two Jacobian matrix results in a quadratic equation for λ. From the form of the solution $(1.4.7)$, the fixed point is stable if for all eigenvalues λ, $\mathrm{Re}\{\lambda\} < 0$, and unstable if for at least one $\lambda, \mathrm{Re}\{\lambda\} > 0$. Here $\mathrm{Re}\{\lambda\}$ means the real part of the (possibly) complex eigenvalue λ. We now reconsider the Lotka-Volterra equations. Fixed point solutions are found by solving $\dot{U} = \dot{V} = 0$, and we have from (1.4.1)

$$U(\alpha - \gamma V) = 0, \quad V(e\gamma U - \beta) = 0$$

Evidently the only two possible solutions are

$$(U_*, V_*) = (0,0) \text{ or } \left(\frac{\beta}{e\gamma}, \frac{\alpha}{\gamma}\right).$$

The trivial fixed point $(0,0)$ is unstable since the prey population grows exponentially if it is initially small. To determine the stability of the second fixed point, we write the Lotka-Volterra equation in the form

$$\frac{dU}{dt} = F(U,V), \quad \frac{dV}{dt} = G(U,V)$$

with

$$F(U,V) = \alpha U - \gamma UV, \quad G(U,V) = e\gamma UV - \beta V$$

The partial derivatives are then computed to be

$$\begin{aligned} F_U &= \alpha - \gamma V, \quad & F_V &= -\gamma U \\ G_U &= e\gamma V, & G_V &= e\gamma U - \beta. \end{aligned}$$

The Jacobian at the fixed point $(U_*, V_*) = (\beta/e\gamma, \alpha/\gamma)$ is

$$J^* = \begin{pmatrix} 0 & -\beta/e \\ e\alpha & 0 \end{pmatrix}$$

and

$$\begin{aligned} \det\left(\mathrm{J}^* - \lambda \mathrm{I}\right) &= \begin{vmatrix} -\lambda & -\beta/e \\ e\alpha & -\lambda \end{vmatrix} \\ &= \lambda^2 + \alpha\beta \\ &= 0 \end{aligned}$$

has the solutions $\lambda_\pm = \pm i\sqrt{\alpha\beta}$, which are pure imaginary. When the eigenvalues of the two-by-two Jacobian are pure imaginary, the fixed point is called a center and the perturbation neither grows nor decays, but oscillates. Here, the angular frequency of oscillation is $\omega = \sqrt{\alpha\beta}$, and the period of the oscillation is $2\pi/\omega$.

We plot U and V versus t (time series plot), and V versus U (phase portraits) to see how the solutions behave. For a nonlinear system of equations such as (1.4.1), a numerical solution is required.

The Lotka-Volterra equations has four free parameters α, β, γ and e. The relevant units here are time, the number of prey, and the number of predators. The Buckingham Pi Theorem predicts that nondimensionalizing the equations can reduce the number of free parameters by three to a manageable single dimensionless grouping of parameters. We choose to nondimensionalize time using the angular frequency of oscillation and the number of prey and predators using their fixed point values. With carets denoting the dimensionless variables, we let

$$\hat{t} = \sqrt{\alpha\beta}t, \quad \hat{U} = U/U_* = \frac{e\gamma}{\beta}U, \quad \hat{V} = V/V_* = \frac{\gamma}{\alpha}V$$

Substitution of (1.4.16) into the Lotka-Volterra equations (1.4.1) results in the dimensionless equations

$$\frac{d\hat{U}}{d\hat{t}} = r(\hat{U} - \hat{U}\hat{V}), \quad \frac{d\hat{V}}{d\hat{t}} = \frac{1}{r}(\hat{U}\hat{V} - \hat{V})$$

with single dimensionless grouping $r = \sqrt{\alpha/\beta}$. Specification of r together with initial conditions completely determines the solution. It should be noted here that the long-time solution of the Lotka-Volterra equations depends on the initial conditions. This asymptotic dependence on the initial conditions is usually considered a flaw of the model.

A numerical solution uses MATLAB's ode45.m built-in function to integrate the differential equations. The code below produces Fig. 1.5. Notice how the predator population lags the prey population: an increase in prey numbers results in a delayed increase in predator numbers as the predators eat more prey. The phase portraits clearly show the periodicity of the oscillation. Note that the curves move counterclockwise: prey numbers increase when predator numbers are minimal, and prey numbers decrease when predator numbers are maximal.

```
function lotka_volterra
% plots time series and phase space diagrams
clear all; close all;
t0=0; tf=6*pi; eps=0.1; delta=0;
r=[1/2, 1, 2];
options = odeset('RelTol',1e-6,'AbsTol',1.e-9);
%time series plots
for i=1:length(r);
   [t,UV]=ode45(@(t,UV) lv_eq(t,UV,r(i)),[t0,tf],[1+eps 1+delta],options);
   U=UV(:,1); V=UV(:,2);
   subplot(3,1,i); plot(t,U,t,V,' --');
   axis([0 6*pi,0.8 1.25]); ylabel('predator,prey');
   text(3,1.15,['r=',num2str(r(i))]);
 end
xlabel('t');
subplot(3,1,1); legend('prey', 'predator');
%phase space plot
xpos=[2.5 2.5 2.5]; ypos=[3.5 3.5 3.5];%for annotating graph
for i=1:length(r);
     for eps=0.1:0.1:1.0;
         [t,UV]=ode45(@(t,UV) lv_eq(t,UV,r(i)),[t0,tf],[1+eps 1+delta],options);
         U=UV(:,1); V=UV(:,2);
         figure(2);subplot(1,3,i); plot(U,V); hold on;
     end
     axis equal; axis([0 4 0 4]);
     text(xpos(i),ypos(i),['r=',num2str(r(i))]);
     if i==1; ylabel('predator'); end;
     xlabel('prey');
end

function dUV=lv_eq(t,UV,r)
dUV=zeros(2,1);
dUV(1) = r*(UV(1)-UV(1)*UV(2));
dUV(2) = (1/r)*(UV(1)*UV(2)-UV(2));
```

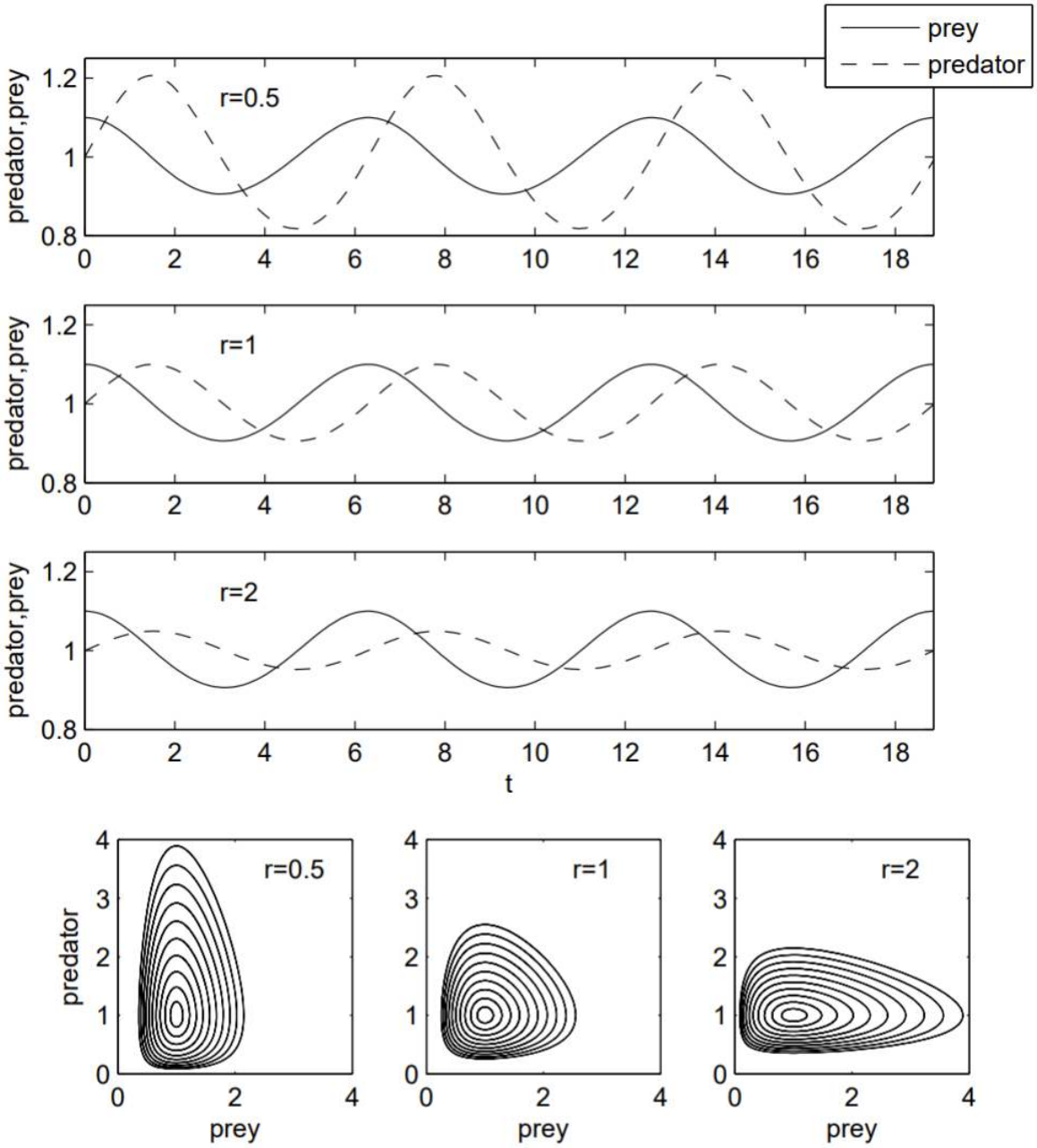


Figure 1.5: Solutions of the dimensionless Lotka-Volterra equations. Upper plots: timeseries solutions; lower plots: phase portraits.
1.4. THE LOTKA-VOLTERRA PREDATOR-PREY MODEL

This page titled 1.4: The Lotka-Volterra Predator-Prey Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

2: Age-structured Populations

Determining the age-structure of a population helps governments plan economic development. Age-structure theory can also help evolutionary biologists better understand a species's life-history. An age-structured population occurs because offspring are born to mothers at different ages. If average per capita birth and death rates at different ages are constant, then a stable age-structure arises. However, a rapid change in birth or death rates can cause the age-structure to shift distributions. In this section, we develop the theory of age-structured populations using both discrete- and continuous-time models. We also present two interesting applications: (1) modeling age-structure changes in China and other countries as these populations age, and; (2) modeling the life cycle of a hermaphroditic worm. We begin this section, however, with one of the oldest problems in mathematical biology: Fibonacci's rabbits. This will lead us to a brief digression about the golden mean, rational approximations and flower development, before returning to our main topic.

2.1: Fibonacci's Rabbits
2.2: The Golden Ratio Φ
2.3: The Fibonacci Numbers in a Sunflower
2.4: Rabbits are an Age-Structured Population
2.5: Discrete Age-Structured Populations
2.6: Continuous Age-Structured Populations
2.7: The Brood Size of a Hermaphroditic Worm

This page titled 2: Age-structured Populations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.1: Fibonacci's Rabbits

In 1202, Fibonacci proposed the following puzzle, which we paraphrase here:

> A man put a male-female pair of newly born rabbits in a field. Rabbits take a month to mature before mating. One month after mating, females give birth to one male-female pair and then mate again. No rabbits die. How many rabbit pairs are there after one year?

The growth of Fibonacci's rabbit population is presented in Table 2.1. At the start of each month, the number of juvenile, adult, and total number of rabbits are shown. At the start of January, one pair of juvenile rabbits is introduced into the population. At the start of February, this pair of rabbits have matured and mate. At the start of March, this original pair of rabbits give birth to a new pair of juvenile rabbits. And so on.

If we let F_n be the total number of rabbit pairs at the start of the nth month, then the number of rabbits pairs at the start of the 13 th month will be the solution to Fibonacci's puzzle. Examining the total number of rabbit pairs in Table 2.1, it is evident that

$$F_{n+1} = F_n + F_{n-1}.$$

This second-order linear difference equation requires two initial conditions, which are given by $F_1 = F_2 = 1$. The first thirteen Fibonacci numbers, read from the table,

Table 2.1: Fibonacci's rabbit population.

month	J	F	M	A	M	J	J	A	S	O	N	D	J
juvenile	1	0	1	1	2	3	5	8	13	21	34	55	89
adult	0	1	1	2	3	5	8	13	21	34	55	89	144
total	1	1	2	3	5	8	13	21	34	55	89	144	233

are given by

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots$$

where $F_{13} = 233$ is the solution to Fibonacci's puzzle.

Let us solve $(2.1.1)$ for all the F_n' s. To solve this equation, we look for a solution of the form $F_n = \lambda^n$. Substitution into (2.1.1) yields

$$\lambda^{n+1} = \lambda^n + \lambda^{n-1}$$

or after division by λ^{n-1} :

$$\lambda^2 - \lambda - 1 = 0$$

with solution

$$\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2}$$

Define

$$\Phi = \frac{1 + \sqrt{5}}{2} = 1.61803\ldots$$

and

$$\phi = \frac{\sqrt{5} - 1}{2} = \Phi - 1 = 0.61803\ldots$$

Then $\lambda_+ = \Phi$ and $\lambda_- = -\phi$. Also, notice that since $\Phi^2 - \Phi - 1 = 0$, division by Φ yields $1/\Phi = \Phi - 1$, so that

$$\phi = \frac{1}{\Phi}$$

As in the solution of linear homogeneous differential equations, the two values of λ can be used to construct a general solution to the linear difference equation using the principle of linear superposition:

$$F_n = c_1\Phi^n + c_2(-\phi)^n.$$

Extending the Fibonacci sequence to $F_0 = 0$ (since $F_0 = F_2 - F_1$), we satisfy the conditions $F_0 = 0$ and $F_1 = 1$:

$$c_1 + c_2 = 0,$$
$$c_1\Phi - c_2\phi = 1.$$

Therefore, $c_2 = -c_1$, and $c_1(\Phi + \phi) = 1$, or $c_1 = 1/\sqrt{5}, c_2 = -1/\sqrt{5}$. We can rewrite the solution as

$$F_n = \frac{\Phi^n - (-\phi)^n}{\sqrt{5}}$$

Since $\phi^n \to 0$ as $n \to \infty$, we see that $F_n \to \Phi^n/\sqrt{5}$, and $F_{n+1}/F_n \to \Phi$

This page titled 2.1: Fibonacci's Rabbits is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.2: The Golden Ratio Φ

The number Φ is known as the golden ratio. Two positive numbers x and y, with $x > y$, are said to be in the golden ratio if the ratio between the sum of those numbers and the larger one is the same as the ratio between the larger one and the smaller; that is,

$$\frac{x+y}{x} = \frac{x}{y}$$

Solution of (2.2.1) yields $x/y = \Phi$. In some well-defined way, Φ can also be called the most irrational of the irrational numbers.

To understand why Φ has this distinction as the most irrational number, we need first to understand continued fractions. Recall that a rational number is any number that can be expressed as the quotient of two integers, and an irrational number is any number that is not rational. Rational numbers have finite continued fractions; irrational numbers have infinite continued fractions.

A finite continued fraction represents a rational number x as

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

where $a_1, a_2, \dots, a_n$ are positive integers and a_0 is any integer. The convenient shorthand form of $(2.2.2)$ is

$$x = [a_0; a_1, a_2, \dots, a_n]$$

If x is irrational, then $n \to \infty$.

Now for some examples. To construct the continued fraction of the rational number $x = 3/5$, we can write

$$\begin{aligned} 3/5 &= \frac{1}{5/3} = \frac{1}{1+2/3} \\ &= \frac{1}{1+\frac{1}{3/2}} = \frac{1}{1+\frac{1}{1+1/2}} \end{aligned}$$

which is of the form $(2.2.2)$, so that $3/5 = [0; 1, 1, 2]$.

To construct the continued fraction of the irrational number $\sqrt{2}$, we can make use of a trick and write

$$\begin{aligned} \sqrt{2} &= 1 + (\sqrt{2} - 1) \\ &= 1 + \frac{1}{1+\sqrt{2}} \end{aligned}$$

We now have a recursive definition that can be continued as

$$\begin{aligned} \sqrt{2} &= 1 + \frac{1}{1 + \left(1 + \frac{1}{1+\sqrt{2}}\right)} \\ &= 1 + \frac{1}{2 + \frac{1}{1+\sqrt{2}}} \end{aligned}$$

and so on, which yields the infinite continued fraction

$$\sqrt{2} = [1; \overline{2}]$$

Another example we will use later is the continued fraction for π, whose first few terms can be calculated from

$$\begin{aligned} \pi &= 3 + 0.14159\dots \\ &= 3 + \frac{1}{7.06251\dots} \\ &= 3 + \frac{1}{7 + \frac{1}{15.99659\dots}} \end{aligned}$$

and so on, yielding the beginning sequence $\pi = [3; 7, 15, \dots]$. The historically important first-order approximation is given by $\pi = [3; 7] = 22/7 = 3.142857\dots$, which was already known by Archimedes in ancient times.

Finally, to determine the continued fraction for the golden ratio Φ, we can write

$$\Phi = 1 + \frac{1}{\Phi}$$

which is another recursive defintion that can be continued as

$$\Phi = 1 + \frac{1}{1 + \frac{1}{\Phi}}$$

and so on, yielding the remarkably simple form

$$\Phi = [1; \overline{1}].$$

Because the trailing a_i 's are all equal to one, the continued fraction for the golden ratio (and other related numbers with trailing ones) converges especially slowly. Furthermore, the successive rational approximations to the golden ratio are just the ratio of consecutive Fibonacci numbers, that is, $1/1, 2/1, 3/2, 5/3$, etc.. Because of the very slow convergence of this sequence, we say that the golden ratio is most difficult to approximate by a rational number. More poetically, the golden ratio has been called the most irrational of the irrational numbers.

Because the golden ratio is the most irrational number, it has a way of appearing unexpectedly in nature. One well-known example is the florets in a sunflower head, which we discuss in the next section.

This page titled 2.2: The Golden Ratio Φ is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.3: The Fibonacci Numbers in a Sunflower

Consider the photo of a sunflower shown in Fig. 2.1, and notice the apparent spirals in the florets radiating out from the center to the edge. These spirals appear to rotate both clockwise and counterclockwise. By counting them, one finds 34 clockwise spirals and 21 counterclockwise spirals. The numbers 21 and 34 are notable as they are consecutive numbers in the Fibonacci sequence.

Why do Fibonacci numbers appear in the sunflower head? To answer this question, we construct a very simple model for the way that the florets develop. Suppose that during development, florets first appear close to the center of the head and subsequently move radially out with constant speed as the sunflower head grows. To fill the circular sunflower head, we suppose that as each new floret is created at the center, it is rotated through a constant angle before moving radially. We will further assume that the angle of rotation is optimum in the sense that the resulting sunflower head consists of florets that are well-spaced.

Figure 2.1: The flowering head of a sunflower.

Let us denote the rotation angle by $2\pi\alpha$. We first consider the possibility that α is a rational number, say n/m, where n and m are positive integers with no common factors, and $n < m$. Since after m rotations florets will return to the radial line on which they started, the resulting sunflower head consists of florets lying along m straight lines. Such a sunflower head for $\alpha = 1/7$ is shown in Fig. 2.2a, where one observes seven straight lines. Evidently, rational values for α do not result in wellspaced florets. This figure and subsequent ones were produced using MATLAB, and the simulation code is shown at the end of this subsection.

What about irrational values? No matter how many rotations, the florets will never return to their starting radial line. Nevertheless, the resulting sunflower head may not have well-spaced florets. For example, if $\alpha = \pi - 3$, then the resulting sunflower head looks like Fig. 2.2b. There, one can see seven counterclockwise spirals. Recall that a good rational approximation to π is $22/7$, which is slightly larger than π. On every seventh counterclockwise rotation, then, the new floret falls just short of the radial line followed by the floret from seven rotations ago.

The irrational numbers that are most likely to construct a sunflower head with well-spaced florets are those that can not be well-approximated by rational numbers. Here, we choose the so-called golden angle, taking $\alpha = 1 - \phi$, so that $2\pi(1 - \phi) \approx 137.5^\circ$, and perform clockwise rotations. The rational approximations to $1 - \phi$ are given by F_n/F_{n+2}, so that the number of spirals observed will correspond to the Fibonacci numbers.

Two simulations of the sunflower head with $\alpha = 1 - \phi$ are shown in Fig. 2.3. These simulations differ only by the choice of radial velocity. In Fig. 2.3a, there can be counted 13 clockwise spirals and 21 counterclockwise spirals; in Fig. 2.3 b, there are 21 counter clockwise spirals and 34 clockwise spirals, just like the sunflower head shown in Fig. 2.1.

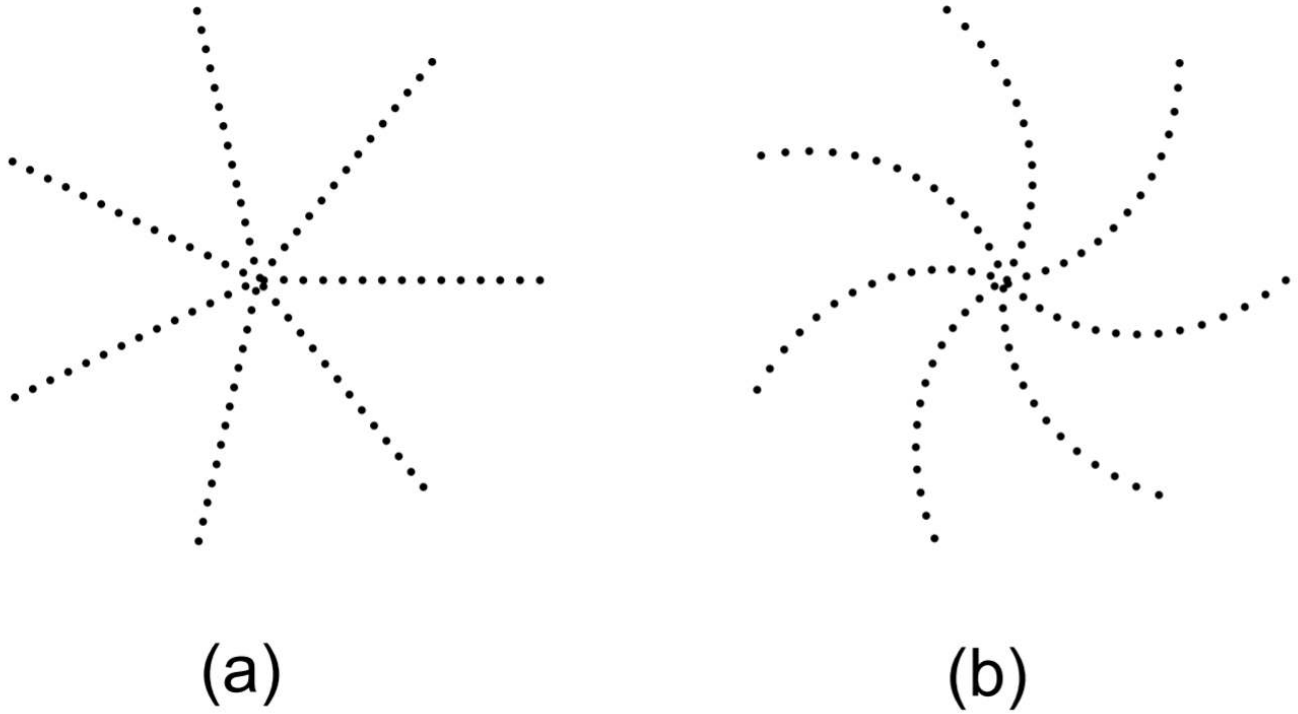


Figure 2.2: Simulation of a sunflower head for (a) $\alpha = 1/7$; (b) $\alpha = \pi - 3$.

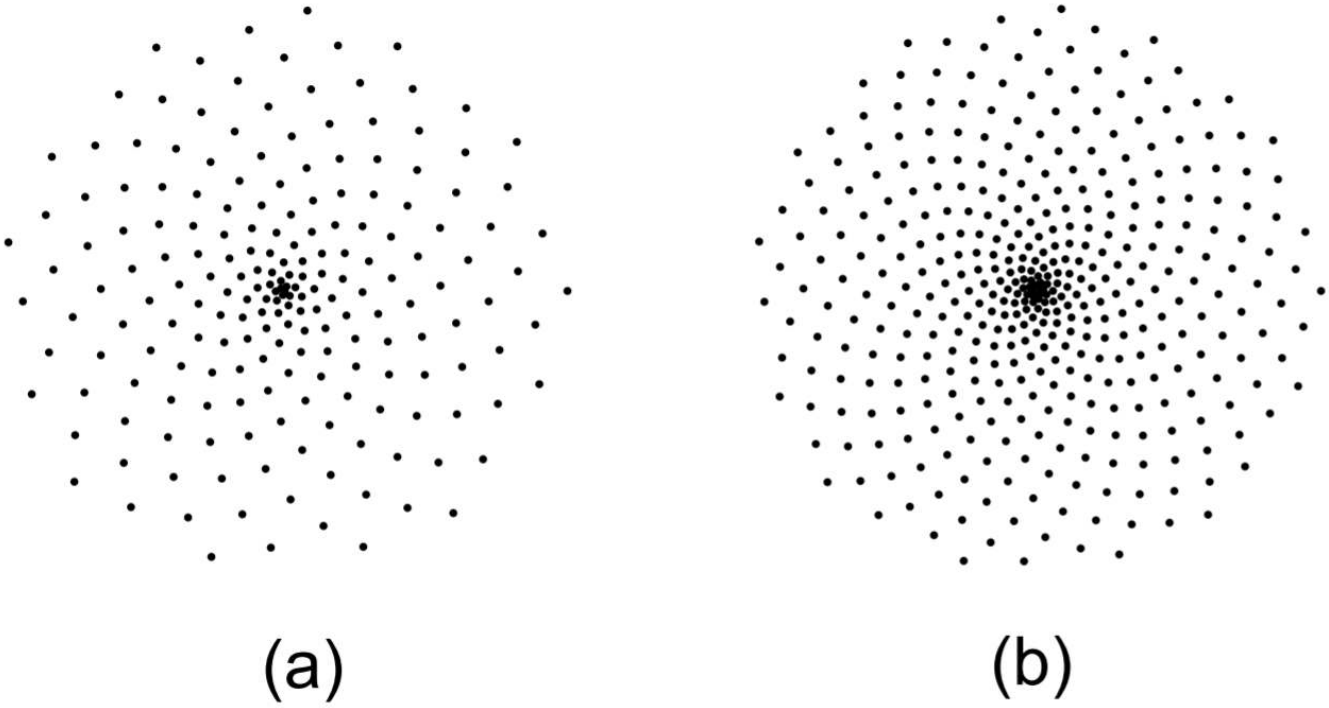


Figure 2.3: Simulation of a sunflower head for $\alpha = 1 - \phi$. (a) $v0 = 1/2$; (b) $v0 = 1/4$.

```
function flower(alpha)
%draws points moving radially but rotated an angle 2*pi*num
figure
angle=2*pi*alpha; r0=0;
v0=1.0; %for higher Fibonacci numbers try v0=0.5, 0.25, 0.1
npts=100/v0;
theta=(0:npts-1)*angle; %all the angles of the npts
r=r0*ones(1,npts); %starting radius of the npts
for i=1:npts % i is total number of points to be plotted (also time)
    for j=1:i %find coordinates of all i points; j=i is newest point
        r(j)=r(j)+v0; %*(i-j);
        x(j)=r(j)*cos(theta(j)); y(j)=r(j)*sin(theta(j));
    end
    plot(x,y,'.','MarkerSize',5.0);
    axis equal;
    axis([-npts*v0-r0 npts*v0+r0 -npts*v0-r0 npts*v0+r0]);
    pause(0.1);
end
%draw a circle at the center
hold on; phi=linspace(0,2*pi); x=r0*cos(phi); y=r0*sin(phi);
fill(x,y,'r')
```

This page titled 2.3: The Fibonacci Numbers in a Sunflower is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.4: Rabbits are an Age-Structured Population

Fibonacci's rabbits form an age-structured population and we can use this simple case to illustrate the more general approach. Fibonacci's rabbits can be categorized into two meaningful age classes: juveniles and adults. Here, juveniles are the newborn rabbits that can not yet mate; adults are those rabbits at least one month old. Beginning with a newborn pair at the beginning of the first month, we census the population at the beginning of each subsequent month after mated females have given birth. At the start of the nth month, let $u_{1,n}$ be the number of newborn rabbit pairs, and let $u_{2,n}$ be the number of rabbit pairs at least one month old. Since each adult pair gives birth to a juvenile pair, the number of juvenile pairs at the start of the $(n+1)$-st month is equal to the number of adult pairs at the start of the n-th month. And since the number of adult pairs at the start of the $(n+1)$-st month is equal to the sum of adult and juvenile pairs at the start of the n-th month, we have

$$u_{1,n+1} = u_{2,n}$$
$$u_{2,n+1} = u_{1,n} + u_{2,n}$$

or written in matrix form

$$\begin{pmatrix} u_{1,n+1} \\ u_{2,n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_{1,n} \\ u_{2,n} \end{pmatrix}.$$

Rewritten in vector form, we have

$$\mathbf{u}_{n+1} = L\mathbf{u}_n,$$

where the definitions of the vector $\mathbf{u}_n$ and the matrix L are obvious. The initial conditions, with one juvenile pair and no adults, are given by

$$\begin{pmatrix} u_{1,1} \\ u_{2,1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The solution of this system of coupled, first-order, linear, difference equations, (2.4.2), proceeds similarly to that of coupled, first-order, linear, differential equations. With the ansatz, $\mathbf{u}_n = \lambda^n \mathbf{v}$, we obtain upon substitution into (2.4.2) the eigenvalue problem

$$\mathbf{L}\mathbf{v} = \lambda\mathbf{v}$$

whose solution yields two eigenvalues λ_1 and λ_2, with corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$. The general solution to (2.4.2) is then

$$\mathbf{u}_n = c_1\lambda_1^n\mathbf{v}_1 + c_2\lambda_2^n\mathbf{v}_2$$

with c_1 and c_2 determined from the initial conditions. Now suppose that $|\lambda_1| > |\lambda_2|$. If we rewrite (2.4.5) in the form

$$\mathbf{u}_n = \lambda_1^n \left(c_1\mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^n \mathbf{v}_2 \right)$$

then because $|\lambda_2/\lambda_1| < 1, \mathbf{u}_n \to c_1\lambda_1^n\mathbf{v}_1$ as $n \to \infty$. The long-time asymptotics of the population, therefore, depends only on λ_1 and the corresponding eigenvector $\mathbf{v}_1$. For our Fibonacci's rabbits, the eigenvalues are obtained by solving $\det(\mathrm{L} - \lambda\mathrm{I}) = 0$, and we find

$$\det \begin{pmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = -\lambda(1-\lambda) - 1$$
$$= 0$$

or $\lambda^2 - \lambda - 1 = 0$, with solutions Φ and $-\phi$. Since $\Phi > \phi$, the eigenvalue Φ and its corresponding eigenvector determine the long-time asymptotic population age-structure. The eigenvector may be found by solving

$$(\mathrm{L} - \Phi\mathrm{I})\mathbf{v}_1 = \mathbf{0}$$

or

$$\begin{pmatrix} -\Phi & 1 \\ 1 & 1-\Phi \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The first equation is just $-\Phi$ times the second equation (use $\Phi^2 - \Phi - 1 = 0$), so that $v_{12} = \Phi v_{11}$. Taking $v_{11} = 1$, we have

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ \Phi \end{pmatrix}$$

The asymptotic age-structure obtained from $\mathbf{v}_1$ shows that the ratio of adults to juveniles approaches the golden mean; that is,

$$\begin{aligned} \lim_{n\to\infty} \frac{u_{2,n}}{u_{1,n}} &= v_{12}/v_{11} \\ &= \Phi. \end{aligned}$$

This page titled 2.4: Rabbits are an Age-Structured Population is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.5: Discrete Age-Structured Populations

In a discrete model, population censuses occur at discrete times and individuals are assigned to age classes, spanning a range of ages. For model simplicity, we assume that the time between censuses is equal to the age span of all age classes.

Table 2.2: Definitions needed in an age-structured, discrete-time population model

$u_{i,n}$	number of females in age class i at census n
s_i	fraction of females surviving from age class $i-1$ to i
m_i	expected number of female offspring from a female in age class i
$l_i = s_1 \cdots s_i$	fraction of females surviving from birth to age class i
$f_i = m_i l_i$	basic reproductive ratio
$\mathcal{R}_0 = \sum_i f_i$	

An example is a country that censuses its population every five years, and assigns individuals to age classes spanning five years (e.g., 0-4 years old, 5-9 years old, etc.). Although country censuses commonly count both females and males separately, we will only count females and ignore males.

There are several new definitions in this Section and I place these in Table 2.2 for easy reference. We define $u_{i,n}$ to be the number of females in age class i at census n. We assume that $i = 1$ represents the first age class and $i = \omega$ the last. No female survives past the last age class. We also assume that the first census takes place when $n = 1$. We define s_i as the fraction of females that survive from age class $i-1$ to age class i (with s_1 the fraction of newborns that survive to their first census), and define m_i as the expected number of female births per female in age class i.

We construct difference equations for $\{u_{in+1}\}$ in terms of $\{u_{in}\}$. First, newborns at census $n+1$ were born between census n and $n+1$ to different aged females, with differing fertilities. Also, only a faction of these newborns survive to their first census. Second, only a fraction of females in age class i that were counted in census n survive to be counted in age class $i+1$ in census $n+1$. Putting these two ideas together with the appropriately defined parameters, the difference equations for $\{u_{i,n+1}\}$ are determined to be

$$\begin{aligned} u_{1,n+1} &= s_1 \left(m_1 u_{1,n} + m_2 u_{2,n} + \cdots + m_\omega u_{\omega,n}\right) \\ u_{2,n+1} &= s_2 u_{1,n} \\ u_{3,n+1} &= s_3 u_{2,n} \\ &\vdots \\ u_{\omega,n+1} &= s_\omega u_{\omega-1,n} \end{aligned}$$

which can be rewritten as the matrix equation

$$\begin{pmatrix} u_{1,n+1} \\ u_{2,n+1} \\ u_{3,n+1} \\ \vdots \\ u_{\omega,n+1} \end{pmatrix} = \begin{pmatrix} s_1 m_1 & s_1 m_2 & \ldots & s_1 m_{\omega-1} & s_1 m_\omega \\ s_2 & 0 & \ldots & 0 & 0 \\ 0 & s_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & s_\omega & 0 \end{pmatrix} \begin{pmatrix} u_{1,n} \\ u_{2,n} \\ u_{3,n} \\ \vdots \\ u_{\omega,n} \end{pmatrix};$$

or in compact vector form as

$$\mathbf{u}_{n+1} = \mathrm{L}\mathbf{u}_n$$

where L is called the Leslie Matrix.

This system of linear equations can be solved by determining the eigenvalues and associated eigenvectors of the Leslie Matrix. One can solve directly the characteristic equation, $\det(\mathrm{L} - \lambda \mathrm{I}) = 0$, or reduce the system of first-order difference equations (2.5.2) to a single high-order equation for the number of females in the first age class. Following the latter approach, and beginning with the second row of $(2.5.2)$, we have

$$
\begin{aligned}
u_{2,n+1} &= s_2 u_{1,n} \\
u_{3,n+1} &= s_3 u_{2,n} \\
&= s_3 s_2 u_{1,n-1} \\
&\vdots \\
u_{\omega,n+1} &= s_\omega u_{\omega-1,n} \\
&= s_\omega s_{\omega-1} u_{\omega-2,n-1} \\
&\vdots \\
&= s_\omega s_{\omega-1} \cdots s_2 u_{1,n-\omega+2}
\end{aligned}
$$

If we define $l_i = s_1 s_2 \cdots s_i$ to be the fraction of females that survive from birth to age class i, and $f_i = m_i l_i$ to be the number of female offspring expected from a newborn female upon reaching age class i (taking into account that she may not survive to age class i), then the first row of (2.5.2) becomes

$$u_{1,n+1} = f_1 u_{1,n} + f_2 u_{1,n-1} + f_3 u_{1,n-2} + \cdots + f_\omega u_{1,n-\omega+1}.$$

Here, we have made the simplifying assumption that $n \geq \omega$ so that all the females counted in the $n+1$ census were born after the first census.

The high-order linear difference equation (2.5.3) may be solved using the ansatz $u_{1,n} = \lambda^n$. Direct substitution and division by λ^{n+1} results in the discrete EulerLotka equation

$$\sum_{j=1}^{\omega} f_j \lambda^{-j} = 1$$

which may have both real and complex-conjugate roots.

Once an eigenvalue λ is determined from (2.5.4), the corresponding eigenvector $\mathbf{v}$ can be computed using the Leslie matrix. We have

$$
\begin{pmatrix}
s_1 m_1 - \lambda & s_1 m_2 & \ldots & s_1 m_{\omega-1} & s_1 m_\omega \\
s_2 & -\lambda & \ldots & 0 & 0 \\
0 & s_3 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & s_\omega & -\lambda
\end{pmatrix}
\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_\omega \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}
$$

Taking $v_\omega = l_\omega / \lambda^\omega$, and beginning with the last row and working backwards, we have:

$$
\begin{aligned}
v_{\omega-1} &= l_{\omega-1} / \lambda^{\omega-1} \\
v_{\omega-2} &= l_{\omega-2} / \lambda^{\omega-2} \\
&\vdots \\
v_1 &= l_1 / \lambda
\end{aligned}
$$

so that

$$v_i = l_i / \lambda^i, \quad \text{for } i = 1, 2, \ldots, \omega$$

We can obtain an interesting implication of this result by forming the ratio of two consecutive age classes. If λ is the dominant eigenvalue (and is real and positive, as is the case for human populations), then asymptotically,

$$
\begin{aligned}
u_{i+1,n} / u_{i,n} &\sim v_{i+1} / v_i \\
&= s_{i+1} / \lambda
\end{aligned}
$$

With the survival fractions $\{s_i\}$ fixed, increasing λ implies a decreasing ratio: a faster growing population has relatively more younger people than a slower growing population. In fact, we are now living through a time when developed countries, particularly Japan and those in Western Europe, as well as Hong Kong and Singapore, have substantially lowered their population growth rates and are increasing the average age of their citizens.

If we want to simply determine if a population grows or decays, we can calculate the basic reproduction ratio $\mathcal{R}_0$, defined as the net expectation of female offspring to a newborn female. Stasis is obtained if the female only replaces herself before dying. If $\mathcal{R}_0 > 1$, then the population grows, and if $\mathcal{R}_0 < 1$ then the population decays. $\mathcal{R}_0$ is equal to the number of female offspring expected from a newborn when she is in age class i, summed over all age classes, or

$$\mathcal{R}_0 = \sum_{i=1}^{\omega} f_i$$

For a population with approximately equal numbers of males and females, $\mathcal{R}_0 = 1$ means a newborn female must produce on average two children over her lifetime. News stories in the western press frequently state that for zero population growth, women need to have 2.1 children. The term women used in these stories presumably means women of child-bearing age. Since girls who die young have no children, the statistic of 2.1 children implies that $0.1/2.1$, or about 5% of children die before reaching adulthood.

A useful application of the mathematical model developed in this Section is to predict the future age structure within various countries. This can be important for economic planning-for instance, determining the tax revenues that can pay for the rising costs of health care as a population ages. For accurate predictions on the future age-structure of a given country, immigration and migration must also be modeled. An interesting website to browse is at

http://www.census.gov/ipc/www/idb.

This website, created by the US census bureau, provides access to the International Data Base (IDB), a computerized source of demographic and socioeconomic statistics for 227 countries and areas of the world. In class, we will look at and discuss the dynamic output of some of the population pyramids, including those for Hong Kong and China.

This page titled 2.5: Discrete Age-Structured Populations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.6: Continuous Age-Structured Populations

We can derive a continuous-time model by considering the discrete model in the limit as the age span Δa of an age class (also equal to the time between censuses) goes to zero. For $n > \omega$, $(2.5.3)$ can be rewritten as

$$u_{1,n} = \sum_{i=1}^{\omega} f_i u_{1,n-i}$$

The first age class in the discrete model consists of females born between two consecutive censuses. The corresponding function in the continuous model is the female birth rate of the population as a whole, $B(t)$, satisfying

$$u_{1,n} = B\left(t_n\right) \Delta a$$

If we assume that the nth census takes place at a time $t_n = n\Delta a$, we also have

$$\begin{aligned} u_{1,n-i} &= B\left(t_{n-i}\right) \Delta a \\ &= B\left(t_n - t_i\right) \Delta a \end{aligned}$$

To determine the continuous analogue of the parameter $f_i = m_i l_i$, we define the agespecific survival function $l(a)$ to be the fraction of newborn females that survive to age a, and define the age-specific maternity function $m(a)$, multiplied by Δa, to be the average number of females born to a female between the ages of a and $a + \Delta a$. With the definition of the age-specific net maternity function, $f(a) = m(a)l(a)$, and $a_i = i\Delta a$, we have

$$f_i = f\left(a_i\right) \Delta a$$

With these new definitions, $(2.6.1)$ becomes

$$B\left(t_n\right) \Delta a = \sum_{i=1}^{\omega} f\left(a_i\right) B\left(t_n - t_i\right) \left(\Delta a\right)^2$$

Canceling one factor of Δa, and using $t_i = a_i$, the right-hand side becomes a Riemann sum. Taking $t_n = t$ and assigning $f(a) = 0$ when a is greater than the maximum age of female fertility, the limit $\Delta a \to 0$ transforms $(2.6.1)$ to

$$B(t) = \int_0^{\infty} B(t-a) f(a) da.$$

Equation (2.6.5) states that the population-wide female birth rate at time t has contributions from females of all ages, and that the contribution to this birth rate from females between the ages of a and $a + da$ is determined from the population-wide female birth rate at the earlier time $t - a$ times the fraction of females that survive to age a times the number of female births to females between the ages of a and $a + da$. Equation $(2.6.5)$ is a linear homogeneous integral equation, valid for t greater than the maximum age of female fertility. A more complete but inhomogeneous equation valid for smaller t can also be derived.

Equation (2.6.5) can be solved by the ansatz $B(t) = e^{rt}$. Direct substitution yields

$$e^{rt} = \int_0^{\infty} f(a) e^{r(t-a)} da$$

which upon canceling e^{rt} results in the continuous Euler-Lotka equation

$$\int_0^{\infty} f(a) e^{-ra} da = 1$$

Equation (2.6.7) is an integral equation for r given the age-specific net maternity function $f(a)$. It is possible to prove that for $f(a)$ a continuous non-negative function, (2.6.7) has exactly one real root r_*, and that the population grows $(r_* > 0)$ or decays $(r_* < 0)$ asymptotically as $e^{r_* t}$. The population growth rate r_* has been called the intrinsic rate of increase, the intrinsic growth rate, or the Malthusian parameter. Typically, (2.6.7) is solved numerically for r using a root-finding algorithm such as Newton's method.

After asymptotically attaining a stable age structure, the population grows like $e^{r_* t}$, and our previous discussion of the Malthusian growth model suggests that r_* may be found from the constant per capita birth rate b and death rate d. By determining expressions for b and d, we will indeed show that $r_* = b - d$.

Figure 2.4: Caenorhabditis elegans, a nematode worm used by biologists as a simple animal model of a multicellular organism. (Photo by Amy Pasquinelli.)

and the difference between the per capita birth and death rates is calculated from

$$b(t) - d(t) = \frac{\int_0^\infty B(t-a)[f(a) - g(a)]da}{\int_0^\infty B(t-a)l(a)da}$$

Asymptotically, a stable age structure is established and the population-wide birth rate grows as $B(t) \sim e^{r_* t}$. Substitution of this expression for $B(t)$ into $(2.6.13)$ and cancelation of $e^{r_* t}$ results in

$$\begin{aligned} b - d &= \frac{\int_0^\infty [f(a) - g(a)]e^{-r_* a}da}{\int_0^\infty l(a)e^{-r_* a}da} \\ &= \frac{1 + \int_0^\infty l'(a)e^{-r_* a}da}{\int_0^\infty l(a)e^{-r_* a}da} \end{aligned}$$

where use has been made of $(2.6.7)$ and $(2.6.11)$. Simplifying the numerator using integration by parts,

$$\begin{aligned} \int_0^\infty l'(a)e^{-r_* a}da &= l(a)e^{-r_* a}\big|_0^\infty + r_* \int_0^\infty l(a)e^{-r_* a}da \\ &= -1 + r_* \int_0^\infty l(a)e^{-r_* a}da \end{aligned}$$

produces the desired result,

$$r_* = b - d$$

It is usually supposed that evolution by natural selection will result in populations with the largest value of the Malthusian parameter r_*, and that natural selection would favor those females that constitute such a population. We will exploit this idea in the next section to compute the brood size of the self-fertilizing hermaphroditic worm of the species Caenorhabditis elegans.

This page titled 2.6: Continuous Age-Structured Populations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

2.7: The Brood Size of a Hermaphroditic Worm

Caenorhabditis elegans, a soil-dwelling nematode worm about $1\ \mathrm{mm}$ in length, is a widely studied model organism in biology. With a body made up of approximately

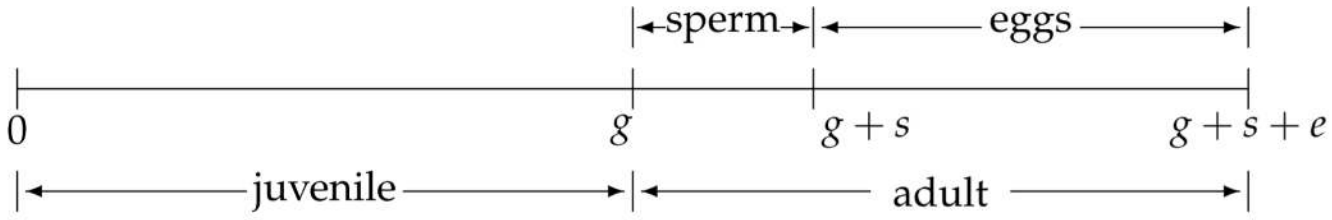


Figure 2.5: A simplified timeline of a hermaphrodite's life.

1000 cells, it is one of the simplest multicellular organisms under study. Advances in understanding the development of this multicellular organism led to the awarding of the 2002 Nobel prize in Physiology or Medicine to the three C. elegans biologists Sydney Brenner, H. Robert Horvitz and John E. Sulston.

The worm C. elegans has two sexes: hermaphrodites, which are essentially females that can produce internal sperm and self-fertilize their own eggs, and males, which must mate with hermaphrodites to produce offspring. In laboratory cultures, males are rare and worms generally propagate by self-fertilization. Typically, a hermaphrodite lays about $250 - 350$ self-fertilized eggs before becoming infertile. It is reasonable to assume that the forces of natural selection have shaped the lifehistory of C. elegans, and that the number of offspring produced by a selfing hermaphrodite must be in some sense optimal. Here, we show how an age-structured model applied to C. elegans yields theoretical insights into the brood size of a selfing hermaphrodite.

To develop a mathematical model for C. elegans, we need to know some details of its life history. As a first approximation (Barker, 1992), a simplified timeline of a hermaphrodite's life is shown in Fig. 2.5. The fertilized egg is laid at time $t = 0$. During a juvenile growth period, the immature worm develops through four larval stages (L1-L4). Towards the end of L4 and for a short while after its final molt to adulthood, the hermaphrodite produces sperm, which is then stored for later use. Then the hermaphrodite produces eggs, self-fertilizes them using her internally stored sperm, and lays them. In the absence of males, egg production ceases after all the sperm are utilized. We assume that the juvenile growth period occurs during $0 < t < g$, spermatogenesis occurs during $g < t < g + s$, and egg-production, selffertilization, and egg laying occurs during $g + s < t < g + s + e$.

Here, we want to understand why hermaphrodites limit their sperm production. Biologists define males and females from the size and metabolic cost of their gametes: sperm are small and cheap and eggs are large and expensive. So on first look, it is puzzling why the total number of offspring produced by a hermaphrodite is limited by the number of sperm produced, rather than by the number of eggs. There must be a hidden cost to the hermaphrodite of producing additional sperm other than metabolic. To understand the basic biology, it is instructive to consider two limiting cases: (1) no sperm production; (2) infinite sperm production. In both cases, the hermaphrodite produces no offspring-in the first case because there are no sperm, and in the second case because there are no eggs. The number of sperm produced by a hermaphrodite before laying eggs is therefore a compromise; although more sperm means more offspring, more sperm also means delayed egg production.

Our main theoretical assumption is that natural selection will favor worms with the ability to establish populations with the largest Malthusian parameter r. Worms containing a genetic mutation resulting in a larger value for r will eventually out-

g	growth period	$72\ \mathrm{h}$
s	sperm production period	$11.9\ \mathrm{h}$
e	egg production period	$65\ \mathrm{h}$
p	sperm production rate	$24\ \mathrm{h}^{-1}$
m	egg production rate	$4.4\ \mathrm{h}^{-1}$
B	brood size	286

Table 2.3: Parameters in a life-history model of C. elegans, with experimental estimates.

number all other worms.

The parameters we need for our mathematical model are listed in Table 2.3, together with estimated experimental values (Cutter, 2004). In addition to the growth period g, sperm production period s, and egg production period e (all in units of hours), we need the sperm production rate p and the egg production rate m (both in units of inverse hours). We also define the brood size B as the total number of fertilized eggs laid by a selfing hermaphrodite. The brood size is equal to the number of sperm produced, and also equal to the number of eggs laid, so that

$$B = ps = me$$

We may use (2.7.1) to eliminate s and e in favor of B :

$$s = B/p, \quad e = B/m$$

The continuous Euler-Lotka equation (2.6.7) for r requires a model for $f(a) = m(a)l(a)$, where $m(a)$ is the age-specific maternity function and $l(a)$ is the agespecific survival function. The function $l(a)$ satisfies the differential equation (2.6.11), and here we make the simplifying assumption that the age-specific mortality function $\mu(a) = d$, where d is the age-independent per capita death rate. Implicitly, we are assuming that worms do not die of old age during egg-laying, but rather die of predation, starvation, disease, or other age-independent causes. Such an assumption is reasonable since worms can live in the laboratory for several weeks after sperm depletion. Solving $(2.6.11)$ with the initial condition $l(0) = 1$ results in

$$l(a) = \exp(-d \cdot a)$$

The age-specific maternity function $m(a)$ is defined such that $m(a)\Delta a$ is the expected number of offspring produced over the age interval Δa. We assume that a hermaphrodite lays eggs at a constant rate m over the ages $g + s < a < g + s + e$; therefore,

$$m(a) = \begin{cases} m & \text{for } g + s < a < g + s + e \\ 0 & \text{otherwise} \end{cases}$$

Using $(2.7.2), (2.7.3)$ and $(2.7.4)$, the continuous Euler-Lotka equation (2.6.7) for the Malthusian parameter r becomes

$$\int_{g+B/p}^{g+B/p+B/m} m \exp[-(r + d)a]da = 1$$

Integrating,

$$\begin{aligned} 1 &= \int_{g+B/p}^{g+B/p+B/m} m \exp[-(r + d)a]da \\ &= \frac{m}{r + d}\{\exp[-(g + B/p)(r + d)] - \exp[-(g + B/p + B/m)(r + d)]\} \\ &= \frac{m}{r + d}\exp[-(g + B/p)(r + d)]\{1 - \exp[-(B/m)(r + d)]\} \end{aligned}$$

which may be rewritten as

$$(r + d) \exp[(g + B/p)(r + d)] = m\{1 - \exp[-(B/m)(r + d)]\}$$

With the parameters d, g, p, and m fixed, the integrated Euler-Lotka equation (2.7.6) is an implicit equation for $r = r(B)$.

To demonstrate that $r = r(B)$ has a maximum at some value of B, we numerically solve (2.7.6) for r with the parameter values of g, p and m obtained from Table 2.3. Since $r + d$ is maximum at the same value of B that r is maximum, and d only enters (2.7.6) in the form $r + d$, without loss of generality we can take $d = 0$. To solve (2.7.6), it is best to make use of Newton's method. We let

$$F(r) = (r + d) \exp[(g + B/p)(r + d)] - m\{1 - \exp[-(B/m)(r + d)]\}$$

and differentiate with respect to r to obtain

$$F'(r) = [1 + (g + B/p)(r + d)] \exp[(g + B/p)(r + d)] - \frac{B}{m}\exp[-(B/m)(r + d)].$$

For a given B, we then solve $F(r) = 0$ by iterating

$$r_{n+1} = r_n - \frac{F(r_n)}{F'(r_n)}$$

Using appropriate starting values for r, the function $r = r(B)$ can be computed and is presented in Fig. 2.6. Evidently, r is maximum near the value $B = 152$, which is 53% of the experimental value for B shown in Table 2.3.

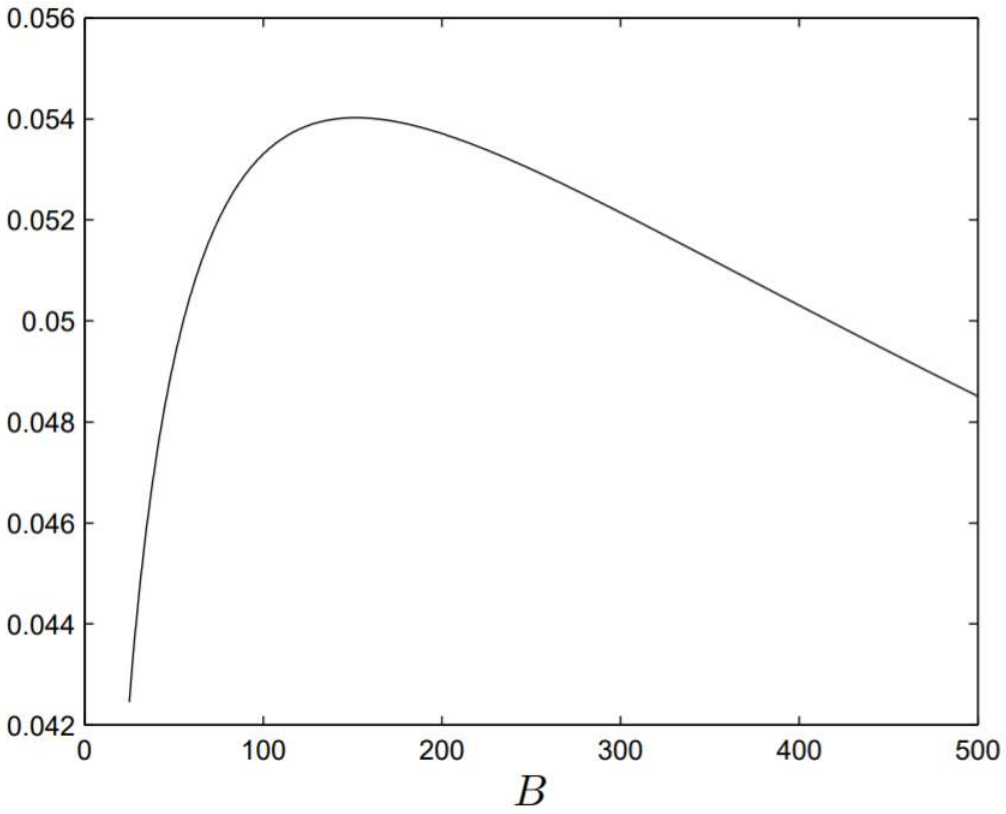


Figure 2.6: A plot of $r = r(B)$, which demonstrates that the Malthusian growth rate r is maximum near a brood size of $B = 152$

Equation (2.7.13) contains the four parameters p, m, g and B, which can be further reduced to three parameters by dimensional analysis. The brood size B is already dimensionless. The rate m at which eggs are produced and laid may be multiplied by the sperm production period $s = B/p$ to form the dimensionless parameter $x = mB/p$. The parameter x represents the number of eggs forgone because of the adult sperm production period and is a measure of the cost of producing sperm. Similarly, m may be multiplied by the larval growth period g to form the dimensionless parameter $y = mg$. The parameter y represents the number of eggs forgone because of the juvenile growth period and is a measure of the cost of development. With B, x and y as our three dimensionless parameters, $(2.7.13)$ becomes

$$\frac{1}{B}\left(1+\frac{B}{x}\right)^{\frac{x+y}{B}} \ln\left(1+\frac{B}{x}\right) = \frac{B}{B+x}.$$

Given values for two of the three dimensionless parameters x, y and $B, (2.7.14)$ may be solved for the remaining parameter, either explicitly for the case of $y = y(x, B)$, or by Newton's method.

The values of x and y obtained from Table 2.3 are $x = 52.5$ and $y = 317$. With $B = 286$, the solution $y = y(x)$ is shown in Fig. 2.7, with the experimental value of (x, y) plotted as a cross.

The seemingly large disagreement between the theoretical result and the experimental data leads us to question the underlying assumptions of the model. Indeed, Cutter (2004) first suggested that the sperm produced precociously as a juvenile does not delay egg production and should be considered cost free. One possibility is to fix the absolute number of sperm produced precociously and to optimize the number of sperm produced as an adult. Another possibility is to fix the fraction of sperm produced precociously and to optimize the total number of sperm produced. This latter assumption was made by Cutter (2004) and seems to best improve the model agreement with the experimental data.

We therefore split the total sperm production period s into juvenile and adult

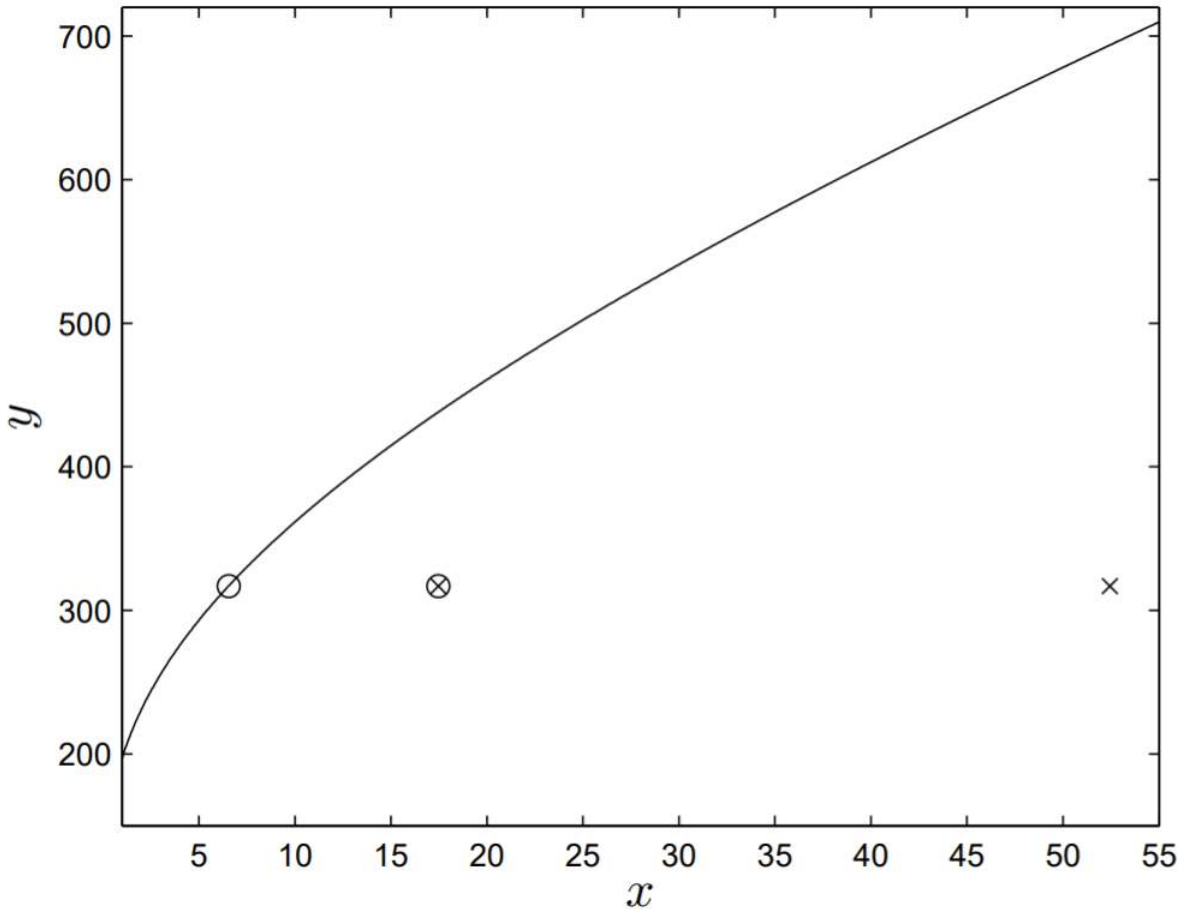


Figure 2.7: Solution curve of y versus x with brood size $B = 286$, obtained by solving (2.7.14). Values for B, m, p and g are taken from Table 2.3. The cross, crossed open circle and open circle correspond to $y = mg$ and $x = mfB/p$ with $f = 1, 1/3$ and $1/8$, respectively.

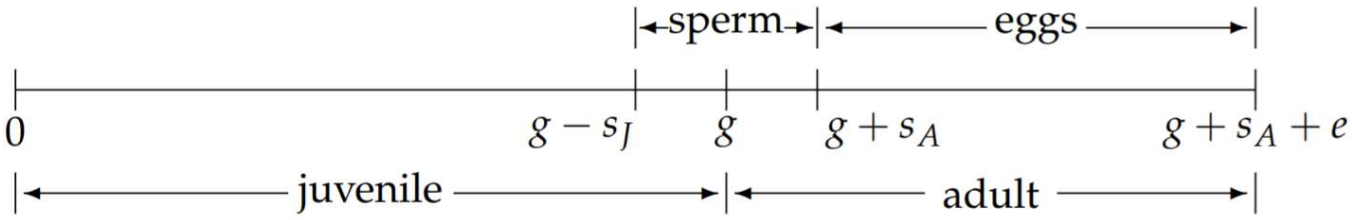


Figure 2.8: A more refined timeline of a hermaphrodite's life.

sperm production periods s_J and s_A, with $s = s_j + s_A$. The revised timeline of the hermaphrodite's life is now shown in Fig. 2.8. With the fraction of sperm produced as an adult denoted by f, and the fraction produced as a juvenile by $1 - f$, we have

$$s_J = (1 - f)s, \quad s_A = fs.$$

Equation (2.7.4) for the age-specific maternity function becomes

$$m(a) = \begin{cases} m & \text{for } g + s_A < a < g + s_A + e \\ 0 & \text{otherwise} \end{cases}$$

with

$$s_A = fB/p.$$

The Euler-Lotka equation (2.7.5) is then changed by the substitution $p \to p/f$. Following this substitution to the final result given by (2.7.14) shows that this equation still holds (which is in fact equivalent to (12) of Chasnov (2011)), but with the now changed definition

$$x = mfB/p$$

$x = mfD/p$

A close examination of the results shown in Fig. 2.7 demonstrates that near-perfect agreement can be made between the results of the theoretical model and the experimental data if $f = 1/8$ (shown as the open circle in Fig. 2.7). Cutter (2004) suggested the value of $f = 1/3$, and this result is shown as a crossed open circle in Fig. 2.7, still in much better agreement with the experimental data than the open circle corresponding to $f = 1$. The added modeling of precocious sperm production through the parameter f thus seems to improve the verisimilitude of the model to the underlying biology.

This page titled 2.7: The Brood Size of a Hermaphroditic Worm is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

3: Stochastic Population Growth

Our derivation of the Malthusian growth model implicitly assumed a large population size. Smaller populations exhibit stochastic effects and these can considerably complicate the modeling. Since in general, modeling stochastic processes in biology is an important yet difficult topic, we will spend some time here analyzing the simplest model of births in finite populations.

3.1: A Stochastic Model of Population Growth

3.2: Asymptotics of Large Initial Populations

3.3: Simulation of Population Growth

This page titled 3: Stochastic Population Growth is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

3.1: A Stochastic Model of Population Growth

The size of the population N is now considered to be a discrete random variable. We define the time-dependent probability mass function $p_N(t)$ of N to be the probability that the population is of size N at time t. Since N must take on one of the values from zero to infinity, we have

$$\sum_{N=0}^{\infty} p_N(t) = 1$$

for all $t \geq 0$. Again, let b be the average per capita birth rate. We make the simplifying approximations that all births are singlets, and that the probability of an individual giving birth is independent of past birthing history. We can then interpret b probabilistically by supposing that as $\Delta t \to 0$, the probability that an individual gives birth during the time Δt is given by $b\Delta t$. For example, if the average per capita birthrate is one offspring every 365 -day year, then the probability that a given individual gives birth on a given day is $1/365$. As we will be considering the limit as $\Delta t \to 0$, we neglect probabilities of more than one birth in the population in the time interval Δt since they are of order $(\Delta t)^2$ or higher. Furthermore, we will suppose that at $t = 0$, the population size is known to be N_0, so that $p_{N_0}(0) = 1$, with all other p_N' s at $t = 0$ equal to zero.

We can determine a system of differential equations for the probability mass function $p_N(t)$ as follows. For a population to be of size $N > 0$ at a time $t + \Delta t$, either it was of size $N - 1$ at time t and one birth occurred, or it was of size N at time t and there were no births; that is

$$p_N(t + \Delta t) = p_{N-1}(t)b(N - 1)\Delta t + p_N(t)(1 - bN\Delta t)$$

Subtracting $p_N(t)$ from both sides, dividing by Δt, and taking the limit $\Delta t \to 0$ results in the forward Kolmogorov differential equations,

$$\frac{dp_N}{dt} = b\left[(N - 1)p_{N-1} - Np_N\right], \quad N = 1, 2, ...$$

where $p_0(t) = p_0(0)$ since a population of zero size remains zero. This system of coupled, first-order, linear differential equations can be solved iteratively. We first review how to solve a first-order linear differential equation of the form

$$\frac{dy}{dt} + ay = g(t), \quad y(0) = y_0$$

where $y = y(t)$ and a is constant. First, we look for an integrating factor μ such that

$$\frac{d}{dt}(\mu y) = \mu\left(\frac{dy}{dt} + ay\right)$$

Differentiating the left-hand-side and multiplying out the right-hand-side results in

$$\frac{d\mu}{dt}y + \mu\frac{dy}{dt} = \mu\frac{dy}{dt} + a\mu y$$

and canceling terms yields

$$\frac{d\mu}{dt} = a\mu.$$

We may integrate this equation with an arbitrary initial condition, so for simplicity we take $\mu(0) = 1$. Therefore, $\mu(t) = e^{at}$. Hence,

$$\frac{d}{dt}\left(e^{at}y\right) = e^{at}g(t)$$

Integrating this equation from 0 to t yields

$$e^{at}y(t) - y(0) = \int_0^t e^{as}g(s)ds$$

Therefore, the solution is

$$y(t) = e^{-at}\left(y(0) + \int_0^t e^{as}g(s)ds\right)$$

The forward Kolmogorov differential equation (3.1.3) is of the form (3.1.4) with $a = bN$ and $g(t) = b(N-1)p_{N-1}$. With the population size known to be N_0 at $t = 0$, the initial conditions can be written succinctly as $p_N(0) = \delta_{N,N_0}$, where δ_{ij} is the Kronecker delta, defined as

$$\delta_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

Therefore, formal integration of (3.1.3) using (3.1.10) results in

$$p_N(t) = e^{-bNt}\left[\delta_{N,N_0} + b(N-1)\int_0^t e^{bNs} p_{N-1}(s)ds\right]$$

The first few solutions of (3.1.12) can now be obtained by successive integrations:

$$p_N(t) = \begin{cases} 0, & \text{if } N < N_0 \\ e^{-bN_0 t}, & \text{if } N = N_0; \\ N_0 e^{-bN_0 t}\left[1 - e^{-bt}\right], & \text{if } N = N_0 + 1; \\ \frac{1}{2}N_0\left(N_0+1\right)e^{-bN_0 t}\left[1-e^{-bt}\right]^2, & \text{if } N = N_0 + 2; \\ \cdots, & \text{if } \ldots \end{cases}$$

Although we will not need this, for completeness I give the complete solution. By defining the binomial coefficient as the number of ways one can select k objects from a set of n identical objects, where the order of selection is immaterial, we have

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

(read as " n choose k''). The general solution for $p_N(t), N \geq N_0$, is known to be

$$p_N(t) = \binom{N-1}{N_0-1} e^{-bN_0 t}\left[1-e^{-bt}\right]^{N-N_0}$$

which statisticians call a shifted negative binomial distribution. The determination of the time-evolution of the probability mass function of N completely solves this stochastic problem.

Of usual main interest is the mean and variance of the population size, and although both could in principle be computed from the probability mass function, we will compute them directly from the differential equation for p_N. The definitions of the mean population size $\langle N\rangle$ and its variance σ^2 are

$$\langle N\rangle = \sum_{N=0}^{\infty} N p_N, \quad \sigma^2 = \sum_{N=0}^{\infty}(N - \langle N\rangle)^2 p_N$$

and we will make use of the equality

$$\sigma^2 = \left\langle N^2\right\rangle - \langle N\rangle^2$$

Multiplying the differential equation (3.1.3) by the constant N, summing over N, and using $p_N = 0$ for $N < N_0$, we obtain

$$\frac{d\langle N\rangle}{dt} = b\left[\sum_{N=N_0+1}^{\infty} N(N-1)p_{N-1} - \sum_{N=N_0}^{\infty} N^2 p_N\right]$$

Now, write $N(N-1) = (N-1)(N-1+1) = (N-1)^2 + (N-1)$, so that the first term on the right-hand-side is

$$\begin{aligned}\sum_{N=N_0+1}^{\infty} N(N-1)p_{N-1} &= \sum_{N=N_0+1}^{\infty}(N-1)^2 p_{N-1} + \sum_{N=N_0+1}^{\infty}(N-1)p_{N-1} \\ &= \sum_{N=N_0}^{\infty} N^2 p_N + \sum_{N=N_0}^{\infty} N p_N\end{aligned}$$

where the second equality was obtained by shifting the summation index downward by one. Therefore, we find the familiar Malthusian growth equation

$$d\langle N\rangle$$

$$\frac{d\langle N\rangle}{dt} = b\langle N\rangle.$$

Together with the initial condition $\langle N\rangle(0) = N_0$, we can find the solution

$$\langle N\rangle(t) = N_0 e^{bt}$$

We proceed similarly to find σ^2 by first determining the differential equation for $\langle N^2\rangle$. Multiplying the differential equation for p_N, (3.1.3), by N^2 and summing over N results in

$$\frac{d\langle N^2\rangle}{dt} = b\left[\sum_{N=N_0+1}^{\infty} N^2(N-1)p_{N-1} - \sum_{N=N_0}^{\infty} N^3 p_N\right]$$

Here, we write $N^2(N-1) = (N-1)(N-1+1)^2 = (N-1)^3 + 2(N-1)^2 + (N-1)$. Proceeding in the same way as above by shifting the index downward, we obtain

$$\frac{d\langle N^2\rangle}{dt} - 2b\langle N^2\rangle = b\langle N\rangle$$

Since $\langle N\rangle$ is known, (3.1.22) is a first-order, linear, inhomogeneous equation for $\langle N^2\rangle$, which can be solved using an integrating factor. The solution obtained using (3.1.10) is

$$\langle N^2\rangle = e^{2bt}\left(N_0^2 + b\int_0^t e^{-2bs}\langle N\rangle(s)ds\right)$$

with $\langle N\rangle$ given by (3.1.20). Performing the integration, we obtain

$$\langle N^2\rangle = e^{2bt}\left[N_0^2 + N_0\left(1-e^{-bt}\right)\right]$$

Finally, using $\sigma^2 = \langle N^2\rangle - \langle N\rangle^2$, we obtain the variance. Thus we arrive at our final result for the population mean and variance:

$$\langle N\rangle = N_0 e^{bt}, \quad \sigma^2 = N_0 e^{2bt}\left(1-e^{-bt}\right)$$

The coefficient of variation c_v measures the standard deviation relative to the mean, and is here given by

$$\begin{aligned} c_v &= \sigma/\langle N\rangle \\ &= \sqrt{\frac{1-e^{-bt}}{N_0}}. \end{aligned}$$

For large t, the coefficient of variation therefore goes like $1/\sqrt{N_0}$, and is small when N_0 is large. In the next section, we will determine the limiting form of the probability distribution for large N_0, recovering both the deterministic model and a Gaussian model approximation.

This page titled 3.1: A Stochastic Model of Population Growth is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

3.2: Asymptotics of Large Initial Populations

Our goal here is to solve for an expansion of the distribution in powers of $1/N_0$ to leading-orders; notice that $1/N_0$ is small if N_0 is large. To zeroth-order, that is in the limit $N_0 \to \infty$, we will show that the deterministic model of population growth is recovered. To first-order in $1/N_0$, we will show that the probability distribution is normal. The latter result will be seen to be a consequence of the well-known Central Limit Theorem in probability theory.

We develop our expansion by working directly with the differential equation for $p_N(t)$. Now, when the population size N is a discrete random variable (taking only the nonnegative integer values of $0, 1, 2, ...$), $p_N(t)$ is the probability mass function for N. If N_0 is large, then the discrete nature of N is inconsequential, and it is preferable to work with a continuous random variable and its probability density function. Accordingly, we define the random variable $x = N/N_0$, and treat x as a continuous random variable, with $0 \le x < \infty$. Now, $p_N(t)$ is the probability that the population is of size N at time t, and the probability density function of x, $P(x,t)$, is defined such that $\int_a^b P(x,t)dx$ is the probability that $a \le x \le b$. The relationship between p and P can be determined by considering how to approximate a discrete probability distribution by a continuous distribution, that is, by defining P such that

$$\begin{aligned} p_N(t) &= \int_{(N-\frac{1}{2})/N_0}^{(N+\frac{1}{2})/N_0} P(x,t)dx \\ &= P\left(N/N_0, t\right)/N_0 \end{aligned}$$

where the last equality becomes exact as $N_0 \to \infty$. Therefore, the appropriate definition for $P(x,t)$ is given by

$$P(x,t) = N_0 p_N(t), \quad x = N/N_0$$

which satisfies

$$\begin{aligned} \int_0^\infty P(x,t)dx &= \sum_{N=0}^{\infty} P\left(N/N_0, t\right)\left(1/N_0\right) \\ &= \sum_{N=0}^{\infty} p_N(t) \\ &= 1 \end{aligned}$$

the first equality (exact only when $N_0 \to \infty$) being a Reimann sum approximation of the integral.

We now transform the infinite set of ordinary differential equations (3.1.3) for $p_N(t)$ into a single partial differential equation for $P(x,t)$. We multiply (3.1.3) by N_0 and substitute $N = N_0 x, p_N(t) = P(x,t)/N_0$, and $p_{N-1}(t) = P\left(x - \frac{1}{N_0}, t\right)/N_0$ to obtain

$$\frac{\partial P(x,t)}{\partial t} = b\left[\left(N_0 x - 1\right) P\left(x - \frac{1}{N_0}, t\right) - N_0 x P(x,t)\right]$$

We next Taylor series expand $P\left(x - 1/N_0, t\right)$ around x, treating $1/N_0$ as a small parameter. That is, we make use of

$$\begin{aligned} P\left(x - \frac{1}{N_0}, t\right) &= P(x,t) - \frac{1}{N_0} P_x(x,t) + \frac{1}{2N_0^2} P_{xx}(x,t) - ... \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n! N_0^n} \frac{\partial^n P}{\partial x^n}. \end{aligned}$$

The two leading terms proportional to N_0 on the right-hand-side of (3.2.2) cancel exactly, and if we group the remaining terms in powers of $1/N_0$, we obtain for the first three leading terms in the expansion

$$\begin{aligned} P_t &= -b\left[\left(xP_x + P\right) - \frac{1}{N_0}\left(\frac{xP_{xx}}{2!} + \frac{P_x}{1!}\right) + \frac{1}{N_0^2}\left(\frac{xP_{xxx}}{3!} + \frac{P_{xx}}{2!}\right) - ...\right] \\ &= -b\left[(xP)_x - \frac{1}{N_0 2!}(xP)_{xx} + \frac{1}{N_0^2 3!}(xP)_{xxx} - ...\right] \end{aligned} \tag{3.2.1}$$

and higher-order terms can be obtained by following the evident pattern.

Equation (3.2.3) may be further analyzed by a perturbation expansion of the probability density function in powers of $1/N_0$:

$$P(x,t) = P^{(0)}(x,t) + \frac{1}{N_0} P^{(1)}(x,t) + \frac{1}{N^2} P^{(2)}(x,t) + ...$$

Here, the unknown functions $P^{(0)}(x,t), P^{(1)}(x,t), P^{(2)}(x,t)$, etc. are to be determined by substituting the expansion (3.2.4) into (3.2.3) and equating coefficients of powers of $1/N_0$. We thus obtain for the coefficients of $(1/N_0)^0$ and $(1/N_0)^1$,

$$P_t^{(0)} = -b\left(xP^{(0)}\right)_x' \tag{3.2.2}$$

$$P_t^{(1)} = -b\left[\left(xP^{(1)}\right)_x - \frac{1}{2}\left(xP^{(0)}\right)_{xx}\right] \tag{3.2.3}$$

Derivation of the deterministic model

The zeroth-order term in the perturbation expansion (3.2.4),

$$P^{(0)}(x,t) = \lim_{N_0\to\infty} P(x,t)$$

satisfies (3.2.5). Equation (3.2.5) is a first-order linear partial differential equation with a variable coefficient. One way to solve this equation is to try the ansatz

$$P^{(0)}(x,t) = h(t)f(r), \quad r = r(x,t),$$

together with the initial condition

$$P^{(0)}(x,0) = f(x)$$

since at $t = 0$, the probability distribution is assumed to be a known function. This initial condition imposes further useful constraints on the functions $h(t)$ and $r(x,t)$:

$$h(0) = 1, \quad r(x,0) = x$$

The partial derivatives of (3.2.8) are

$$P_t^{(0)} = h'f + hr_tf', \quad P_x^{(0)} = hr_xf',$$

which after substitution into (3.2.5) results in

$$\left(h' + bh\right)f + \left(r_t + bxr_x\right)hf' = 0$$

This equation can be satisfied for any f provided that $h(t)$ and $r(x,t)$ satisfy

$$h' + bh = 0, \quad r_t + bxr_x = 0$$

The first equation for $h(t)$, together with the initial condition $h(0) = 1$, is easily solved to yield

$$h(t) = e^{-bt}.$$

To determine a solution for $r(x,t)$, we try the technique of separation of variables. We write $r(x,t) = X(x)T(t)$, and upon substitution into the differential equation for $r(x,t)$, we obtain

$$XT' + bxX'T = 0$$

and division by XT and separation yields

$$\frac{T'}{T} = -bx\frac{X'}{X}$$

Since the left-hand-side is independent of x, and the right-hand-side is independent of t, both the left-hand-side and right-hand-side must be constant, independent of both x and t. Now, our initial condition is $r(x,0) = x$, so that $X(x)T(0) = x$. Without lose of generality, we can take $T(0) = 1$, so that $X(x) = x$. The right-handside of (3.2.16) is therefore equal to the constant $-b$, and we obtain the differential equation and initial condition

$$T' + bT = 0, \quad T(0) = 1$$

which we can solve to yield $T(t) = e^{-bt}$. Therefore, our solution for $r(x,t)$ is

$$r(x,t) = xe^{-bt}$$

By putting our solutions (3.2.14) and (3.2.18) together into our ansatz (3.2.8), we have obtained the general solution to the pde:

$$P^{(0)}(x,t) = e^{-bt} f\left(xe^{-bt}\right)$$

To determine f, we apply the initial condition of the probability mass function, $p_N(0) = \delta_{N,N_0}$. From (3.10), the corresponding initial condition on the probability distribution function is

$$P(x,0) = \begin{cases} N_0 & \text{if } 1 - \frac{1}{2N_0} \leq x \leq 1 + \frac{1}{2N_0} \\ 0 & \text{otherwise} \end{cases}$$

In the limit $N_0 \to \infty, P(x,0) \to P^{(0)}(x,0) = \delta(x-1)$, where $\delta(x-1)$ is the Dirac delta-function, centered around 1 . The delta-function is widely used in quantum physics and was introduced by Dirac for that purpose. It now finds many uses in applied mathematics. It can be defined by requiring that, for any function $g(x)$,

$$\int_{-\infty}^{+\infty} g(x)\delta(x)dx = g(0)$$

The usual view of the delta-function $\delta(x-a)$ is that it is zero everywhere except at $x = a$ at which it is infinite, and its integral is one. It is not really a function, but it is what mathematicians call a distribution.

Now, since $P^{(0)}(x,0) = f(x) = \delta(x-1)$, our solution becomes

$$P^{(0)}(x,t) = e^{-bt}\delta\left(xe^{-bt} - 1\right)$$

This can be rewritten by noting that (letting $y = ax - c$),

$$\begin{aligned} \int_{-\infty}^{+\infty} g(x)\delta(ax-c)dx &= \frac{1}{a}\int_{-\infty}^{+\infty} g((y+c)/a)\delta(y)dy \\ &= \frac{1}{a}g(c/a) \end{aligned}$$

yielding the identity

$$\delta(ax-c) = \frac{1}{a}\delta\left(x - \frac{c}{a}\right)$$

From this, we can rewrite our solution $(3.2.22)$ in the more intuitive form

$$P^{(0)}(x,t) = \delta\left(x - e^{bt}\right)$$

Using $(3.2.23)$, the zeroth-order expected value of x is

$$\begin{aligned} \langle x_0 \rangle &= \int_0^{\infty} xP^{(0)}(x,t)dx \\ &= \int_0^{\infty} x\delta\left(x - e^{bt}\right)dx \\ &= e^{bt} \end{aligned}$$

while the zeroth-order variance is

$$\begin{aligned} \sigma_{x_0}^2 &= \left\langle x_0^2 \right\rangle - \langle x_0 \rangle^2 \\ &= \int_0^{\infty} x^2 P^{(0)}(x,t)dx - e^{2bt} \\ &= \int_0^{\infty} x^2\delta\left(x - e^{bt}\right)dx - e^{2bt} \\ &= e^{2bt} - e^{2bt} \\ &= 0 \end{aligned}$$

Thus, in the infinite population limit, the random variable x has zero variance, and is therefore no longer random, but follows $x = e^{bt}$ deterministically. We say that the probability distribution of x becomes sharp in the limit of large population sizes. The

general principle of modeling large populations deterministically can simplify mathematical models when stochastic effects are unimportant.

Derivation of the normal probability distribution

We now consider the first-order term in the perturbation expansion (3.2.4), which satisfies (3.2.6). We do not know how to solve (3.2.6) directly, so we will attempt to find a solution following a more circuitous route. First, we proceed by computing the moments of the probability distribution. We have

$$\begin{aligned}\langle x^n\rangle &= \int_0^\infty x^n P(x,t)dx \\ &= \int_0^\infty x^n P^{(0)}(x,t)dx + \frac{1}{N_0}\int_0^\infty x^n P^{(1)}(x,t)dx + ... \\ &= \langle x_0^n\rangle + \frac{1}{N_0}\langle x_1^n\rangle + ...,\end{aligned}$$

where the last equality defines $\langle x_0^n\rangle$, etc. Now, using (3.2.23),

$$\begin{aligned}\langle x_0^n\rangle &= \int_0^\infty x^n P^{(0)}(x,t)dx \\ &= \int_0^\infty x^n \delta\left(x-e^{bt}\right)dx \\ &= e^{nbt}\end{aligned} \tag{3.2.4}$$

To determine $\langle x_1^n\rangle$, we use $(3.2.6)$. Multiplying by x^n and integrating, we have

$$\frac{d\langle x_1^n\rangle}{dt} = -b\left[\int_0^\infty x^n\left(xP^{(1)}\right)_x dx - \frac{1}{2}\int_0^\infty x^n\left(xP^{(0)}\right)_{xx}dx\right]$$

We integrate by parts to remove the derivatives of xP, assuming that xP and all its derivatives vanish on the boundaries of integration, where x is equal to zero or infinity. We have

$$\begin{aligned}\int_0^\infty x^n\left(xP^{(1)}\right)_x dx &= -\int_0^\infty nx^n P^{(1)}dx \\ &= -n\langle x_1^n\rangle\end{aligned}$$

and

$$\begin{aligned}\frac{1}{2}\int_0^\infty x^n\left(xP^{(0)}\right)_{xx}dx &= -\frac{n}{2}\int_0^\infty x^{n-1}\left(xP^{(0)}\right)_x dx \\ &= \frac{n(n-1)}{2}\int_0^\infty x^{n-1}P^{(0)}dx \\ &= \frac{n(n-1)}{2}\langle x_0^{n-1}\rangle\end{aligned}$$

Therefore, after integration by parts, (3.2.25) becomes

$$\frac{d\langle x_1^n\rangle}{dt} = -b\left[n\langle x_1^n\rangle + \frac{n(n-1)}{2}\langle x_0^{n-1}\rangle\right]$$

Equation (3.2.26) is a first-order linear inhomogeneous differential equation and can be solved using an integrating factor (see (3.1.10) and the preceding discussion). Solving this differential equation using (3.2.24) and the initial condition $\langle x_1^n\rangle(0)=0$, we obtain

$$\langle x_1^n\rangle = \frac{n(n-1)}{2}e^{nbt}\left(1-e^{-bt}\right)$$

The probability distribution function, accurate to order $1/N_0$, may be obtained by making use of the so-called moment generating function $\Psi(s)$, defined as

$$\begin{aligned}\Psi(s) &= \langle e^{sx}\rangle \\ &= 1 + s\langle x\rangle + \frac{s^2}{}\langle x^2\rangle + \frac{s^3}{}\langle x^3\rangle + ...\end{aligned}$$

$$= \sum_{n=0}^{\infty} \frac{s^n \langle x^n \rangle}{n!}$$

To order $1/N_0$, we have

$$\Psi(s) = \sum_{n=0}^{\infty} \frac{s^n \langle x_0^n \rangle}{n!} + \frac{1}{N_0} \sum_{n=0}^{\infty} \frac{s^n \langle x_1^n \rangle}{n!} + \mathrm{O}\left(1/N_0^2\right)$$

Now, using (3.2.24),

$$\sum_{n=0}^{\infty} \frac{s^n \langle x_0^n \rangle}{n!} = \sum_{n=0}^{\infty} \frac{\left(se^{bt}\right)^n}{n!}$$
$$= e^{se^{bt}}$$

and using (3.2.27),

$$\sum_{n=0}^{\infty} \frac{s^n \langle x_1^n \rangle}{n!} = \frac{1}{2}\left(1-e^{-bt}\right) \sum_{n=0}^{\infty} \frac{1}{n!} n(n-1) \left(se^{bt}\right)^n$$
$$= \frac{1}{2}\left(1-e^{-bt}\right) s^2 e^{2bt} \sum_{n=0}^{\infty} \frac{1}{n!} n(n-1) \left(se^{bt}\right)^{n-2}$$
$$= \frac{1}{2}\left(1-e^{-bt}\right) s^2 \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^2}{\partial s^2} \left(se^{bt}\right)^n$$
$$= \frac{1}{2}\left(1-e^{-bt}\right) s^2 \frac{\partial^2}{\partial s^2} \sum_{n=0}^{\infty} \frac{\left(se^{bt}\right)^n}{n!}$$
$$= \frac{1}{2}\left(1-e^{-bt}\right) s^2 \frac{\partial^2}{\partial s^2} \left(e^{se^{bt}}\right)$$
$$= \frac{1}{2}\left(1-e^{-bt}\right) s^2 e^{2bt} e^{se^{bt}}$$

Therefore,

$$\Psi(s) = e^{se^{bt}} \left(1 + \frac{1}{2N_0}\left(1-e^{-bt}\right) s^2 e^{2bt} + \ldots\right)$$

We can recognize the parenthetical term of (3.2.29) as a Taylor-series expansion of an exponential function truncated to first-order, i.e.,

$$\exp\left(\frac{1}{2N_0}\left(1-e^{-bt}\right) s^2 e^{2bt}\right) = 1 + \frac{1}{2N_0}\left(1-e^{-bt}\right) s^2 e^{2bt} + \mathrm{O}\left(1/N_0^2\right)$$

Therefore, to first-order in $1/N_0$, we have

$$\Psi(s) = \exp\left(se^{bt} + \frac{1}{2N_0}\left(1-e^{-bt}\right) s^2 e^{2bt}\right) + \mathrm{O}\left(1/N_0^2\right)$$

Standard books on probability theory (e.g., A first course in probability by Sheldon Ross, pg. 365) detail the derivation of the moment generating function of a normal random variable:

$$\Psi(s) = \exp\left(\langle x \rangle s + \frac{1}{2}\sigma^2 s^2\right), \quad \text{for a normal random variable;}$$

and comparing (3.2.32) with (3.2.31) shows us that the probability distribution $P(x,t)$ to first-order in $1/N_0$ is normal with the mean and variance given by

$$\langle x \rangle = e^{bt}, \quad \sigma_x^2 = \frac{1}{N_0} e^{2bt}\left(1-e^{-bt}\right)$$

The mean and variance of $x = N/N_0$ is equivalent to those derived for N in (3.1.25), but now we learn that N is approximately normal for large populations.

The appearance of a normal probability distribution (also called a Gaussian probability distribution) in a first-order expansion is in fact a particular case of the Central Limit Theorem, one of the most important and useful theorems in probability and statistics. We state here a simple version of this theorem without proof:

Central Limit Theorem: Suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (iid) random variables with mean $\langle X \rangle$ and variance σ_X^2. Then for sufficiently large n, the probability distribution of the average of the X_i 's, denoted as the random variable $Z = \frac{1}{n}\sum_{i=1}^{n} X_i$, is well approximated by a Gaussian with mean $\langle X \rangle$ and variance $\sigma^2 = \sigma_X^2/n$

The central limit theorem can be applied directly to our problem. Consider that our population consists of N_0 founders. If $m_i(t)$ denotes the number of individuals descendent from founder i at time t (including the still living founder), then the total number of individuals at time t is $N(t) = \sum_{i=1}^{N_0} m_i(t)$; and the average number of descendants of a single founder is $x(t) = N(t)/N_0$. If the mean number of descendants of a single founder is $\langle m \rangle$, with variance σ_m^2, then by applying the central limit theorem for large N_0, the probability distribution function of x is well approximated by a Gaussian with mean $\langle x \rangle = \langle m \rangle$ and variance $\sigma_x^2 = \sigma_m^2/N_0$. Comparing with our results (3.2.33), we find $\langle m \rangle = e^{bt}$ and $\sigma_m^2 = e^{2bt}\left(1 - e^{-bt}\right)$.

This page titled 3.2: Asymptotics of Large Initial Populations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

3.3: Simulation of Population Growth

As we have seen, stochastic modeling is significantly more complicated than deterministic modeling. As the modeling becomes more sophisticated, a numerical simulation becomes necessary. Here, for illustration, we show how to simulate individual realizations of population growth.

A naive approach would make use of the birth rate b directly. During the short time interval Δt, each individual has probability $b\Delta t$ of giving birth. We can decide if an individual gives birth by generating a random deviate (a pseudo-random number between zero and one): if the random deviate is less than $b\Delta t$, then the individual gives birth; if larger than $b\Delta t$, then the individual does not. With N individuals at time t, we then simply compute N random deviates. Counting the number of random deviates less than $b\Delta t$ allows us to update the population size to the time $t + \Delta t$. For accuracy, Δt must be small, making this a computationally slow method.

There is, however, a much more efficient way to simulate population growth. Define a random variable $\tau = \tau(N)$ to be the time it takes for a population to grow from size N to size $N + 1$ because of a single birth. The random variable τ is called the interevent time and represents the elapsed time between births. A simulation from population size N_0 to size N_f would then simply require computing $N_f - N_0$ different random values of τ, a relatively easy and quick computation if we know the probability density function (pdf) of τ.

Accordingly, we define $P(\tau)$ to be the pdf of τ for a population of size N. The cumulative distribution function (cdf), $F(\tau)$, defined as the probability that the interevent time is less that τ is given by

$$F(\tau) = \int_0^\tau P(\tau) d\tau$$

where $P(\tau) = F'(\tau)$. The complementary cumulative distribution function (ccdf), $G(\tau)$, defined as the probability that the interevent time is greater than τ is given by $G(\tau) = 1 - F(\tau)$

Now the probability that the interevent time is greater than $\tau + \Delta\tau$, with $\Delta\tau$ small, is given by the probability that it is greater than τ times the probability that there are no births in the time interval $\Delta\tau$. Therefore, $G(\tau + \Delta\tau)$ satisfies

$$G(\tau + \Delta\tau) = G(\tau)(1 - bN\Delta\tau)$$

Differencing G and taking the limit $\Delta\tau \to 0$ yields the differential equation

$$\frac{dG}{d\tau} = -bNG$$

which can be integrated using the initial condition $G(0) = 1$ to obtain

$$G(\tau) = e^{-bN\tau}$$

From $G(\tau)$, we can find

$$F(\tau) = 1 - e^{-bN\tau}, \quad P(\tau) = bNe^{-bN\tau}$$

The pdf, $P(\tau)$, has the form of an exponential distribution with parameter bN.

Here, we make use of a well-known result from probability theory that enables us to compute τ using random deviates. With y a random deviate, τ can be computed from $\tau = F^{-1}(y)$, where F^{-1} is the inverse function of F. The correct formula for the exponential distribution is

$$\tau = -\frac{\ln(1 - y)}{bN}$$

To simulate a population growing from N_0 to N_f, we compute $N_f - N_0$ random deviates y, and then compute the corresponding interevent times using (3.3.6), taking care to adjust the population size N as the population grows.

Below, we illustrate a simple MATLAB function that simulates one realization of population growth from initial size N_0 to final size N_f, with birth rate b.

```
function [t, N] = population_growth_simulation(b,N0,Nf)
% simulates population growth from N0 to Nf with birth rate b
N=N0:Nf;
y=rand(1,Nf-N0); % random deviates
tau=-log(1-y)./(b*N(1:Nf-N0)); % interevent times
t=[0 cumsum(tau)]; % cumulative sum of interevent times
```

The function population_growth_simulation.m can be driven by a MATLAB script to compute realizations of population growth. For instance, the following script computes 25 realizations for a population growth from 10 to 100 with $b = 1$ and plots all the realizations:

```
% calculate nreal realizations and plot
b=1; N0=10; Nf=100;
nreal=25;
for i=1:nreal
    [t,N]=population_growth_simulation(b,N0,Nf);
    plot(t,N); hold on;
end
xlabel('t'); ylabel('N');
```

Figure 3.1 presents three graphs, showing 25 realizations of population growth starting with population sizes of 10,100 , and 1000 , and ending with population sizes a factor of 10 larger. Observe that the variance, relative to the initial population size, decreases as the initial population size increases, following our analytical result (3.31). (a)

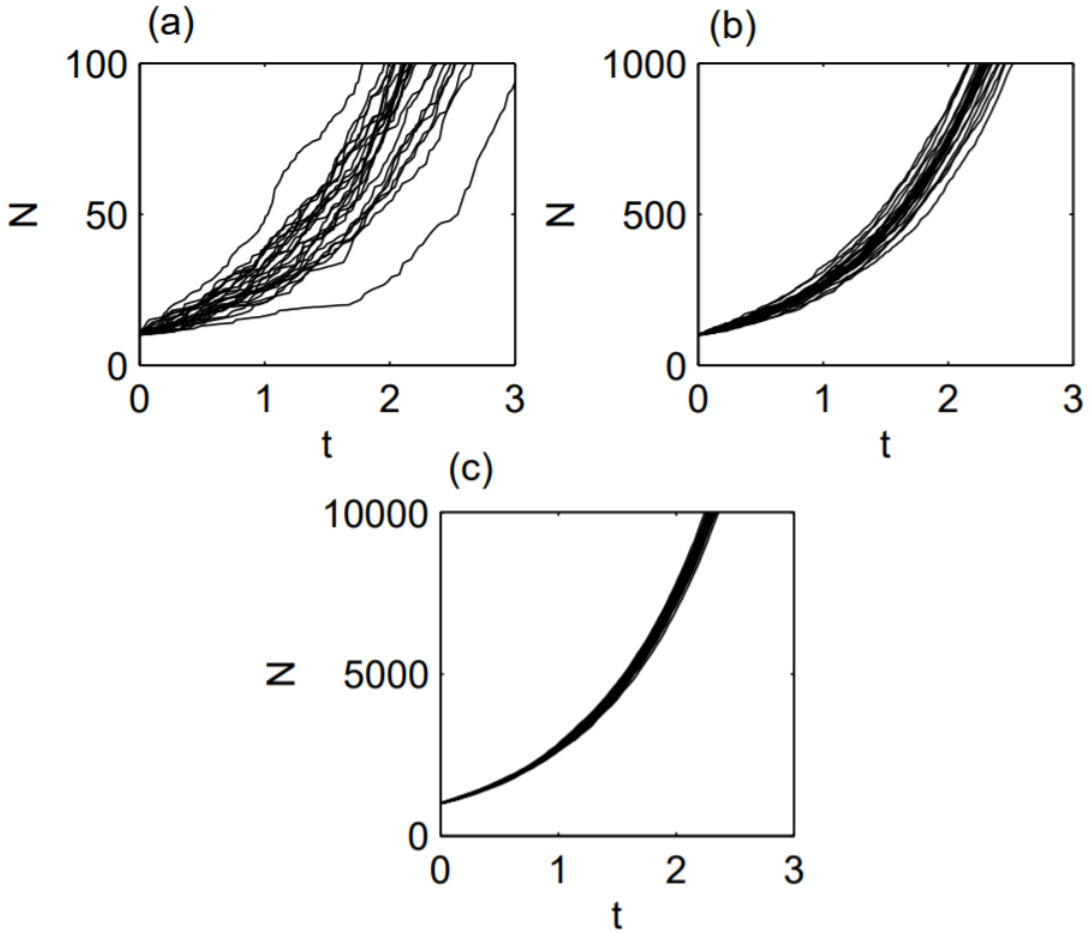


Figure 3.1: Twenty-five realizations of population growth with initial population sizes of 10,100 , and 1000 , in (a), (b), and (c), respectively.

This page titled 3.3: Simulation of Population Growth is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

4: Infectious Disease Modeling

In the late $1320'$ s, an outbreak of the bubonic plague occurred in China. This disease is caused by the bacteria Yersinia pestis and is transmitted from rats to humans by fleas. The outbreak in China spread west, and the first major outbreak in Europe occurred in 1347 . During a five year period, 25 million people in Europe, approximately $1/3$ of the population, died of the black death. Other more recent epidemics include the influenza pandemic known as the Spanish flu killing 50-100 million people worldwide during the years $1918 - 1919$, and the present AIDS epidemic, originating in Africa and first recognized in the USA in 1981 , killing more than 25 million people. For comparison, the SARS epidemic for which Hong Kong was the global epicenter in the year 2003 resulted in 8096 known SARS cases and 774 deaths. Yet, we know well that this relatively small epidemic caused local social and economic turmoil.

Here, we introduce the most basic mathematical models of infectious disease epidemics and endemics. These models form the basis of the necessarily more detailed models currently used by world health organizations, both to predict the future spread of a disease and to develop strategies for containment and eradication.

4.1: The SI Model
4.2: The SIS Model
4.3: The SIR Epidemic Disease Model
4.4: Vaccination
4.5: The SIR Endemic Disease Model
4.6: Evolution of Virulence

This page titled 4: Infectious Disease Modeling is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

4.1: The SI Model

The simplest model of an infectious disease categorizes people as either susceptible or infective (SI). One can imagine that susceptible people are healthy and infective people are sick. A susceptible person can become infective by contact with an infective. Here, and in all subsequent models, we assume that the population under study is well mixed so that every person has equal probability of coming into contact with every other person. This is a major approximation. For example, while the population of Amoy Gardens could be considered well mixed during the SARS epidemic because of shared water pipes and elevators, the population of Hong Kong as a whole could not because of the larger geographical distances, and the limited travel of many people outside the neighborhoods where they live.

We derive the governing differential equation for the SI model by considering the number of people that become infective during time Δt. Let $\beta\Delta t$ be the probability that a random infective person infects a random susceptible person during time Δt. Then with S susceptible and I infective people, the expected number of newly infected people in the total population during time Δt is $\beta\Delta tSI$. Thus,

$$I(t+\Delta t) = I(t) + \beta\Delta t S(t) I(t)$$

and in the limit $\Delta t \to 0$,

$$\frac{dI}{dt} = \beta SI$$

We diagram (4.1.1) as

$$S \overset{\beta SI}{\longrightarrow} I.$$

Later, diagrams will make it easier to construct more complicated systems of equations. We now assume a constant population size N, neglecting births and deaths, so that $S + I = N$. We can eliminate S from (4.1.1) and rewrite the equation as

$$\frac{dI}{dt} = \beta N I \left(1 - \frac{I}{N}\right)$$

which can be recognized as a logistic equation, with growth rate βN and carrying capacity N. Therefore $I \to N$ as $t \to \infty$ and the entire population will become infective.

This page titled 4.1: The SI Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

4.2: The SIS Model

The SI model may be extended to the SIS model, where an infective can recover and become susceptible again. We assume that the probability that an infective recovers during time Δt is given by $\gamma \Delta t$. Then the total number of infective people that recover during time Δt is given by $I \times \gamma \Delta t$, and

$$I(t+\Delta t) = I(t) + \beta \Delta t S(t) I(t) - \gamma \Delta t I(t)$$

or as $\Delta t \to 0$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

which we diagram as

$$S \underset{\gamma I}{\overset{\beta SI}{\rightleftharpoons}} I.$$

Using $S + I = N$, we eliminate S from (4.2.2) to obtain

$$\frac{dI}{dt} = (\beta N - \gamma) I \left(1 - \frac{\beta}{\beta N - \gamma} I\right)$$

which is again a logistic equation, but now with growth rate $\beta N - \gamma$ and carrying capacity $N - \gamma/\beta$. In the SIS model, an epidemic will occur if $\beta N > \gamma$. And if an epidemic does occur, then the disease becomes endemic with the number of infectives at equilibrium given by $I_* = N - \gamma/\beta$, and the number of susceptibles given by $S_* = \gamma/\beta$.

In general, an important metric for whether or not an epidemic will occur is called the basic reproductive ratio. The basic reproductive ratio is defined as the expected number of people that a single infective will infect in an otherwise susceptible population. To compute the basic reproductive ratio, define $l(t)$ to be the probability that an individual initially infected at $t = 0$ is still infective at time t. Since the probability of being infective at time $t + \Delta t$ is equal to the probability of being infective at time t multiplied by the probability of not recovering during time Δt, we have

$$l(t+\Delta t) = l(t)(1 - \gamma \Delta t)$$

or as $\Delta t \to 0$

$$\frac{dl}{dt} = -\gamma l$$

With initial condition $l(0) = 1$

$$l(t) = e^{-\gamma t}$$

Now, the expected number of secondary infections produced by a single primary infective over the time period $(t, t + \Delta t)$ is given by the probability that the primary infective is still infectious at time t multiplied by the expected number of secondary infections produced by a single infective during time Δt; that is, $l(t) \times S(t) \beta \Delta t$. Here, the definition of the basic reproductive ratio assumes that the entire population is susceptible so that $S(t) = N$. Therefore, the expected number of secondary infectives produced by a single primary infective in a completely susceptible population is

$$\begin{aligned} \int_0^\infty \beta l(t) N dt &= \beta N \int_0^\infty e^{-\gamma t} dt \\ &= \frac{\beta N}{\gamma} \end{aligned}$$

The basic reproductive ratio, written as $\mathcal{R}_0$, is therefore defined as

$$\mathcal{R}_0 = \frac{\beta N}{\gamma}$$

and from (4.2.4), we can see that in the SIS model an epidemic will occur if $R_0 > 1$. In other words, an epidemic can occur if an infected individual in an otherwise susceptible population will on average infect more than one other individual.

We have also seen an analogous definition of the basic reproductive ratio in our previous discussion of age-structured populations ($\S 2.5$). There, the basic reproductive ratio was the number of female offspring expected from a new born female over her lifetime; the population size would grow if this value was greater than unity.

In the SIS model, after an epidemic occurs the population reaches an equilibrium between susceptible and infective individuals. The effective basic reproductive ratio of this steady-state population can be defined as $\beta S_*/\gamma$, and with $S_* = \gamma/\beta$ this ratio is evidently unity. Clearly, for a population to be in equilibrium, an infective individual must infect on average one other individual before he or she recovers.

This page titled 4.2: The SIS Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

4.3: The SIR Epidemic Disease Model

The SIR model, first published by Kermack and McKendrick in 1927, is undoubtedly the most famous mathematical model for the spread of an infectious disease. Here, people are characterized into three classes: susceptible S, infective I and removed R. Removed individuals are no longer susceptible nor infective for whatever reason; for example, they have recovered from the disease and are now immune, or they have been vaccinated, or they have been isolated from the rest of the population, or perhaps they have died from the disease. As in the SIS model, we assume that infectives leave the I class with constant rate γ, but in the SIR model they move directly into the R class. The model may be diagrammed as

$$S \xrightarrow{\beta SI} I \xrightarrow{\gamma I} R,$$

and the corresponding coupled differential equations are

$$\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma I,$$

with the constant population constraint $S + I + R = N$. For convenience, we nondimensionalize (4.3.2) using N for population size and γ^{-1} for time; that is, let

$$\hat{S} = S/N, \quad \hat{I} = I/N, \quad \hat{R} = R/N, \quad \hat{t} = \gamma t$$

and define the dimensionless basic reproductive ratio as

$$\mathcal{R}_0 = \frac{\beta N}{\gamma}$$

The dimensionless SIR equations are then given by

$$\frac{d\hat{S}}{d\hat{t}} = -\mathcal{R}_0 \hat{S}\hat{I}, \quad \frac{d\hat{I}}{d\hat{t}} = \mathcal{R}_0 \hat{S}\hat{I} - \hat{I}, \quad \frac{d\hat{R}}{d\hat{t}} = \hat{I},$$

with dimensionless constraint $\hat{S} + \hat{I} + \hat{R} = 1$.

We will use the SIR model to address two fundamental questions: (1) Under what condition does an epidemic occur? (2) If an epidemic occurs, what fraction of a well-mixed population gets sick?

Let $\left(\hat{S}_*, \hat{I}_*, \hat{R}_*\right)$ be the fixed points of (4.3.5). Setting $d\hat{S}/d\hat{t} = d\hat{I}/d\hat{t} = d\hat{R}/d\hat{t} = 0$, we immediately observe from the equation for $d\hat{R}/d\hat{t}$ that $\hat{I} = 0$, and this value forces all the time-derivatives to vanish for any $\hat{S}$ and $\hat{R}$. Since with $\hat{I} = 0$, we have $\hat{R} = 1 - \hat{S}$, evidently all the fixed points of (4.3.5) are given by the one parameter family $\left(\hat{S}_*, \hat{I}_*, \hat{R}_*\right) = \left(\hat{S}_*, 0, 1 - \hat{S}_*\right)$.

An epidemic occurs when a small number of infectives introduced into a susceptible population results in an increasing number of infectives. We can assume an initial population at a fixed point of $(4.3.5)$, perturb this fixed point by introducing a small number of infectives, and determine the fixed point's stability. An epidemic occurs when the fixed point is unstable. The linear stability problem may be solved by considering only the equation for $d\hat{I}/d\hat{t}$ in (4.3.5). With $\hat{I} \ll 1$ and $\hat{S} \approx \hat{S}_0$, we have

$$\frac{d\hat{I}}{d\hat{t}} = \left(\mathcal{R}_0 \hat{S}_0 - 1\right) \hat{I}$$

so that an epidemic occurs if $\mathcal{R}_0 \hat{S}_0 - 1 > 0$. With the basic reproductive ratio given by (4.3.4), and $S_0 = S_0/N$, where S_0 is the number of initial susceptible individuals, an epidemic occurs if

$$\mathcal{R}_0 \hat{S}_0 = \frac{\beta S_0}{\gamma} > 1$$

which could have been guessed. An epidemic occurs if an infective individual introduced into a population of S_0 susceptible individuals infects on average more than one other person. If an epidemic occurs, then initially the number of infective individuals increases exponentially with growth rate $\beta S_0 - \gamma$.

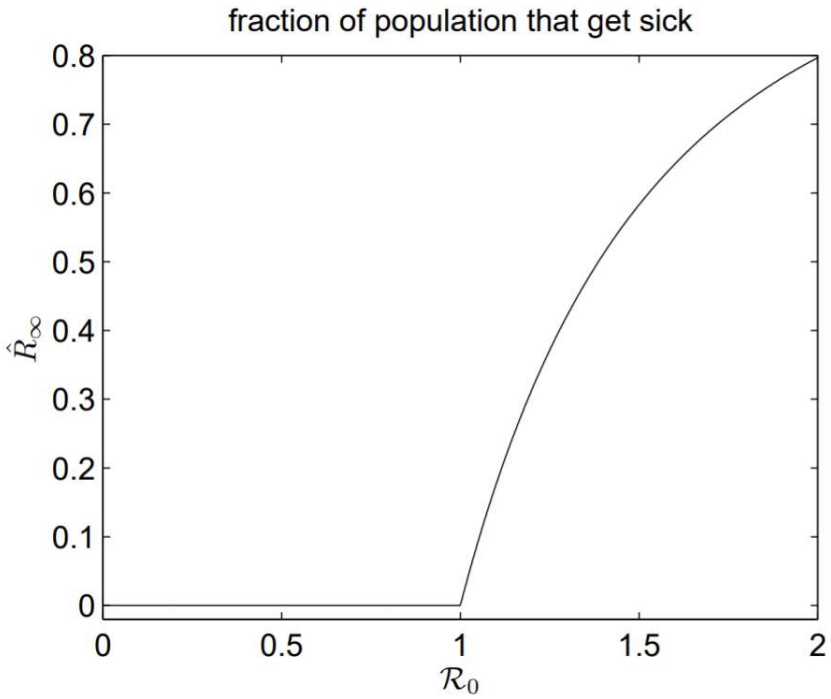


Figure 4.1: The fraction of the population that gets sick in the SIR model as a function of the basic reproduction ratio $\mathcal{R}_0$.

which upon integration and simplification, results in the following transcendental equation for $\hat{R}_\infty$:

$$1 - \hat{R}_\infty - e^{-\mathcal{R}_0 \hat{R}_\infty} = 0,$$

an equation that can be solved numerically using Newton's method. We have

$$F\left(\hat{R}_\infty\right) = 1 - \hat{R}_\infty - e^{-\mathcal{R}_0 \hat{R}_\infty}$$

$$F'\left(\hat{R}_\infty\right) = -1 + \mathcal{R}_0 e^{-\mathcal{R}_0 \hat{R}_\infty}$$

and Newton's method for solving $F\left(\hat{R}_\infty\right) = 0$ iterates

$$\hat{R}_\infty^{(n+1)} = \hat{R}_\infty^n - \frac{F\left(\hat{R}_\infty^{(n)}\right)}{F'\left(\hat{R}_\infty^{(n)}\right)}$$

for fixed $\mathcal{R}_0$ and a suitable initial condition for $R_\infty^{(0)}$, which we take to be unity. My code for computing R_∞ as a function of $\mathcal{R}_0$ is given below, and the result is shown in Fig. 4.1. There is an explosion in the number of infections as $\mathcal{R}_0$ increases from unity, and this rapid increase is a classic example of what is known more generally as a threshold phenomenon.

```
function [R0, R_inf] = sir_rinf
% computes solution of R_inf using Newton's method from SIR model
nmax=10; numpts=1000;
R0 = linspace(0,2,numpts); R_inf = ones(1,numpts);
for i=1:nmax
    R_inf = R_inf - F(R_inf,R0)./Fp(R_inf,R0);
end
plot(R0,R_inf); axis([0 2 -0.02 0.8])
xlabel('$\mathcal{R}_0$', 'Interpreter', 'latex', 'FontSize',16)

ylabel('$\hat R_\infty$', 'Interpreter', 'latex', 'FontSize',16);
title('fraction of population that get sick')
%subfunctions
function y = F(R_inf,R0)
y = 1 - R_inf - exp(-R0.*R_inf);
function y = Fp(R_inf,R0)
y = -1 + R0.*exp(-R0.*R_inf);
```

This page titled 4.3: The SIR Epidemic Disease Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

4.4: Vaccination

Table 4.1 lists the diseases for which vaccines exist and are widely administered to children. Health care authorities must determine the fraction of a population that must be vaccinated to prevent epidemics.

We address this problem within the SIR epidemic disease model. Let p be the fraction of the population that is vaccinated and p_* the minimum fraction required to prevent an epidemic. When $p > p_*$, an epidemic can not occur. Since even non-vaccinated people are protected by the absence of epidemics, we say that the population has acquired herd immunity.

We assume that individuals are susceptible unless vaccinated, and vaccinated individuals are in the removed class. The initial population is then modeled as $(\hat{S}, \hat{I}, \hat{R}) = (1-p, 0, p)$. We have already determined the stability of this fixed point to perturbation by a small number of infectives. The condition for an epidemic to occur is given by (4.3.7), and with $\hat{S}_0 = 1-p$, an epidemic occurs if

$$\mathcal{R}_0(1-p) > 1$$

Therefore, the minimum fraction of the population that must be vaccinated to prevent an epidemic is

$$p_* = 1 - \frac{1}{\mathcal{R}_0}$$

Diseases with smaller values of $\mathcal{R}_0$ are easier to eradicate than diseases with larger values $\mathcal{R}_0$ since a population can acquire herd immunity with a smaller fraction of the population vaccinated. For example, smallpox with $\mathcal{R}_0 \approx 4$ has been eradicated throughout the world whereas measles with $\mathcal{R}_0 \approx 17$ still has occasional outbreaks.

This page titled 4.4: Vaccination is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

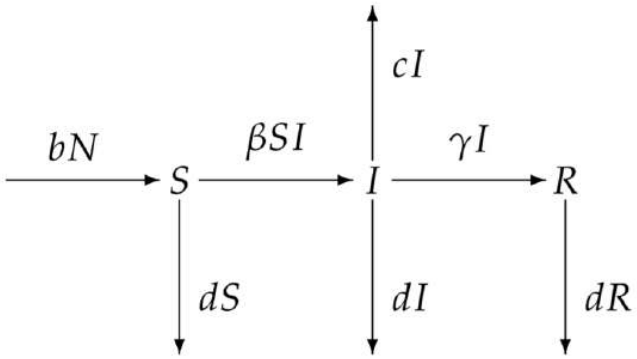
4.5: The SIR Endemic Disease Model

A disease that is constantly present in a population is said to be endemic. For example, malaria is endemic to Sub-Saharan Africa, where about 90% of malaria related deaths occur. Endemic diseases prevail over long time scales: babies are born, old people die.

Table 4.1: Previously common diseases for which vaccines have been developed.

Disease	Description	Symptoms	Complications
Diphtheria	A bacterial respiratory disease	Sore throat and lowgrade fever	Airway obstruction, coma, and death
Haemophilus influenzae type b (Hib)	A bacterial infection occurring primarily in infants	Skin and throat infections, meningitis, pneumonia, sepsis, and arthritis	Death in one out of 20 children, and permanent brain damage in 10% - 30% of the survivors
Hepatitis A	A viral liver disease	Potentially none; yellow skin or eyes, tiredness, stomach ache, loss of appetite, or nausea	usually none
Hepatitis B	Same as Hepatitis A	Same as Hepatitis A	Life-long liver problems, such as scarring of the liver and liver cancer
Measles	A viral respiratory disease	Rash, high fever, cough, runny nose, and red, watery eyes	Diarrhea, ear infections, pneumonia, encephalitis, seizures, and death
Mumps	A viral lymph node disease	Fever, headache, muscle ache, and swelling of the lymph nodes close to the jaw	Meningitis, inflammation of the testicles or ovaries, inflammation of the pancreas and deafness
Pertussis (whooping cough)	A bacterial respiratory disease	Severe spasms of coughing	Pneumonia, encephalitis, and death, especially in infants
Pneumococcal disease	A bacterial disease	High fever, cough, and stabbing chest pains, bacteremia, and meningitis	Death
Polio	A viral lymphatic and nervous system disease	Fever, sore throat, nausea, headaches, stomach aches, stiffness in the neck, back, and legs	Paralysis that can lead to permanent disability and death
Rubella (German measles)	A viral respiratory disease	Rash and fever for two to three days	Birth defects if acquired by a pregnant woman
Tetanus (lockjaw)	A bacterial nervous system disease	Lockjaw, stiffness in the neck and abdomen, and difficulty swallowing	Death in one third of the cases, especially people over age 50
Varicella (chickenpox)	A viral disease in the Herpes family	A skin rash of blister-like lesions	Bacterial infection of the skin, swelling of the brain, and pneumonia
Human papillomavirus	A viral skin and mucous membrane disease	Warts, cervical cancer	The 5-year survival rate from all diagnoses of cervical cancer is 72%

Let b be the birth rate and d the disease-unrelated death rate. We separately define c to be the disease-related death rate; R is now the immune class. We may diagram a SIR model of an endemic disease as



and the governing differential equations are

$$\frac{dS}{dt} = bN - \beta SI - dS, \quad \frac{dI}{dt} = \beta SI - (d + c + \gamma)I, \quad \frac{dR}{dt} = \gamma I - dR$$

with $N = S + I + R$. In our endemic disease model, N separately satisfies the differential equation

$$dN/dt = (b - d)N - cI$$

and is not necessarily constant.

A disease can become endemic in a population if dI/dt stays nonnegative; that is,

$$\frac{\beta S(t)}{d + c + \gamma} \geq 1$$

For a disease to become endemic, newborns must introduce an endless supply of new susceptibles into a population.

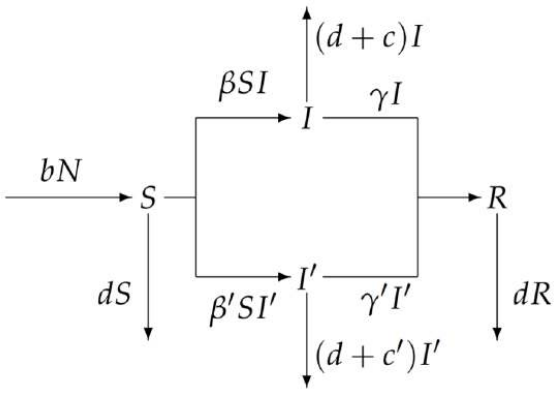
This page titled 4.5: The SIR Endemic Disease Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

4.6: Evolution of Virulence

Microorganisms continuously evolve due to selection pressures in their environments. Antibiotics are a common source of selection pressure on pathogenic bacteria, and the development of antibiotic-resistant strains presents a major health challenge to medical science. Bacteria and viruses also compete directly with each other for reproductive success resulting in the evolution of virulence. Here, using the SIR endemic disease model, we study how virulence may evolve.

For the sake of argument, we will assume that a population is initially in equilibrium with an endemic disease caused by a wildtype virus; that is, S, I and R are assumed to be nonzero and at equilibrium values. Now suppose that some virus particles mutate by a random, undirected process that occurs naturally. We want to determine the conditions under which the mutant virus will replace the wildtype virus in the population. In mathematical terms, we want to determine the linear stability of the endemic disease equilibrium to the introduction of a mutant viral strain.

We assume that the original wildtype virus has infection rate β, removal rate γ, and disease-related death rate c, and that the mutant virus has corresponding rates β', γ' and c'. We further assume that an individual infected with either a wildtype or mutant virus gains immunity to subsequent infection from both wildtype and mutant viral forms. Our model thus has a single susceptible class S, two distinct infective classes I and I' depending on which virus causes the infection, and a single recovered class R. The appropriate diagram is



with corresponding differential equations

$$\frac{dS}{dt} = bN - dS - S(\beta I + \beta' I') \tag{4.6.1}$$

$$\frac{dI}{dt} = \beta SI - (d + c + \gamma)I \tag{4.6.2}$$

$$\frac{dI'}{dt} = \beta' SI' - (d + c' + \gamma')I' \tag{4.6.3}$$

$$\frac{dR}{dt} = \gamma I + \gamma' I' - dR \tag{4.6.4}$$

If the population is initially in equilibrium with the wildtype virus, then we have $\dot{I} = 0$ with $I \neq 0$, and the equilibrium value for S is determined from (4.6.2) to be

$$S_* = \frac{d + c + \gamma}{\beta},$$

which corresponds to a basic reproductive ratio $\beta S_*/(d + c + \gamma)$ of unity.

We perturb this endemic disease equilibrium by introducing a small number of infectives carrying the mutated virus, that is, by letting I' be small. Rather than solve the stability problem by means of a Jacobian analysis, we can directly examine the equation for dI'/dt given by (4.6.3). Here, with $S = S_*$ given by (4.6.5), we have

$$\frac{dI'}{dt} = \left[\frac{\beta'(d + c + \gamma)}{\beta} - (d + c' + \gamma')\right] I'$$

and I' increases exponentially if

$$\frac{\beta'(d+c+\gamma)}{\beta} - (d+c'+\gamma') > 0$$

or after some elementary algebra,

$$\frac{\beta'}{d+c'+\gamma'} > \frac{\beta}{d+c+\gamma}$$

Our result (4.6.8) suggests that endemic viruses (or other microorganisms) will tend to evolve (i) to be more easily transmitted between people $(\beta' > \beta)$; (ii) to make people sick longer $(\gamma' < \gamma)$, and; (iii) to be less deadly $c' < c$. In other words, viruses evolve to increase their basic reproductive ratios. For instance, our model suggests that viruses evolve to be less deadly because the dead do not spread disease. Our result would not be applicable, however, if the dead in fact did spread disease, a possibility if disposal of the dead was not done with sufficient care, perhaps because of certain cultural traditions such as family washing of the dead body.

This page titled 4.6: Evolution of Virulence is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

5: Population Genetics

Deoxyribonucleic acid, or DNA-a large double-stranded, helical molecule, with rungs made from the four base pairs adenine (A), cytosine (C), thymine (T) and guanine (G)-carries inherited genetic information. The ordering of the base pairs A, C, T and G determines the DNA sequence. A gene is a particular DNA sequence that is the fundamental unit of heredity for a particular trait. Some species develop as diploids, carrying two copies of every gene, one from each parent, and some species develop as haploids with only one copy. There are even species that develop as both diploids and haploids.

Consider the pea plant, which develops as a diploid. When we say there is a gene for pea color, say, we mean there is a particular DNA sequence that may vary in a pea plant population, and that there are at least two subtypes, called alleles, where plants with two copies of the yellow-color allele have yellow peas, those with two copies of the green-color allele, green peas. A plant with two copies of the same allele is homozygous for that particular gene (or a homozygote), while a plant carrying two different alleles is heterozygous (or a heterozygote). For the pea color gene, a plant carrying both a yellow- and green-color allele has yellow peas. We say that the green color is a recessive trait (or the green-color allele is recessive), and the yellow color is a dominant trait (or the yellow-color allele is dominant). The combination of alleles carried by the plant is called its genotype, while the actual trait (green or yellow peas) is called its phenotype. A gene that has more than one allele in a population is called polymorphic, and we say the population has a polymorphism for that particular gene.

Population genetics can be defined as the mathematical modeling of the evolution and maintenance of polymorphism in a population. Population genetics together with Charles Darwin's theory of evolution by natural selection and Gregor Mendel's theory of biological inheritance forms the modern evolutionary synthesis (sometimes called the modern synthesis, the evolutionary synthesis, the neoDarwinian synthesis, or neo-Darwinism). The primary founders in the early twentieth century of population genetics were Sewall Wright, J. B. S. Haldane and Ronald Fisher.

Allele frequencies in a population can change due to the influence of four primary evolutionary forces: natural selection, genetic drift, mutation, and migration. Here, we mainly focus on natural selection and mutation. Genetic drift is the study of stochastic effects, and it is important in small populations. Migration typically requires consideration of the spatial distribution of a population, and it is usually modeled mathematically by partial differential equations.

The simplified models we will consider assume infinite population sizes (neglecting stochastic effects except in $\$5.5$), well-mixed populations (neglecting any spatial distribution), and discrete generations (neglecting any age-structure). Our main purpose is to illustrate the fundamental ways that a genetic polymorphism can be maintained in a population.

Table 5.1: Haploid genetics using population size, absolute viability, and fertility fitnesses.

genotype	A	a
number	n_A	n_a
viability fitness	g_A	g_a
fertility fitness	f_A	f_a

5.1: Haploid Genetics
5.2: Diploid Genetics
5.3: Frequency-Dependent Selection
5.4: Linkage Equilibrium
5.5: Random Genetic Drift

This page titled 5: Population Genetics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

5.1: Haploid Genetics

We first consider the modeling of selection in a population of haploid organisms. Selection is modeled by fitness coefficients, with different genotypes having different fitnesses. We begin with a simple model that counts the number of individuals in the next generation, and then show how this model can be reformulated in terms of allele frequencies and relative fitness coefficients.

Table 5.1 formulates the basic model. We assume that there are two alleles A and a for a particular haploid gene. These alleles are carried in the population by n_A and n_a individuals, respectively. A fraction $g_A\,(g_a)$ of individuals carrying allele $A(a)$ is assumed to survive to reproduction age, and those that survive contribute $f_A\,(f_a)$ offspring to the next generation. These are of course average values, but under the assumption of an (almost) infinite population, our model is deterministic. Accordingly, with $n_A^{(i)}\left(n_a^{(i)}\right)$ representing the number of individuals carrying allele $A(a)$ in the i th generation, and formulating a discrete generation model, we have

$$n_A^{(i+1)} = f_A g_A n_A^{(i)}, \quad n_a^{(i+1)} = f_a g_a n_a^{(i)}$$

It is mathematically easier and more transparent to work with allele frequencies rather than individual numbers. We denote the frequency (or more accurately, proportion) of allele $A(a)$ in the i th generation by $p_i\,(q_i)$; that is,

$$p_i = \frac{n_A^{(i)}}{n_A^{(i)} + n_a^{(i)}}, \quad q_i = \frac{n_a^{(i)}}{n_A^{(i)} + n_a^{(i)}}$$

where evidently $p_i + q_i = 1$. Now, from (5.1.1),

$$n_A^{(i+1)} + n_a^{(i+1)} = f_A g_A n_A^{(i)} + f_a g_a n_a^{(i)},$$

so that dividing the first equation in (5.1.1) by (5.1.3) yields

$$\begin{aligned} p_{i+1} &= \frac{f_A g_A n_A^{(i)}}{f_A g_A n_A^{(i)} + f_a g_a n_a^{(i)}} \\ &= \frac{f_A g_A p_i}{f_A g_A p_i + f_a g_a q_i} \\ &= \frac{\left(\frac{f_A g_A}{f_a g_a}\right) p_i}{\left(\frac{f_A g_A}{f_a g_a}\right) p_i + q_i} \end{aligned} \tag{5.1.1}$$

where the second equality comes from dividing the numerator and denominator by $n_A^{(i)} + n_a^{(i)}$, and the third equality from dividing the numerator and denominator by

genotype	A	a
freq. of gamete	p	q
relative fitness	$1+s$	1
freq after selection	$(1+s)p/w$	q/w
normalization	$w=(1+s)p+q$	

Table 5.2: Haploid genetic model of the spread of a favored allele.

$f_A g_A$. Similarly,

$$q_{i+1} = \frac{q_i}{\left(\frac{f_A g_A}{f_a g_a}\right) p_i + q_i}$$

which could also be derived using $q_{i+1} = 1 - p_{i+1}$. We observe from the evolution equations for the allele frequencies, (5.1.4) and (5.1.5), that only the relative fitness $f_A g_A / f_a g_a$ of the alleles matters. Accordingly, in our models, we will consider only relative fitnesses, and we will arbitrarily set one fitness to unity to simplify the algebra and make the final result more transparent.

Spread of a favored allele

We consider a simple model for the spread of a favored allele in Table 5.2, with $s > 0$. Denoting p' by the frequency of A in the next generation (not (!) the derivative of p), the evolution equation is given by

$$\begin{aligned} p' &= \frac{(1+s)p}{w} \\ &= \frac{(1+s)p}{1+sp} \end{aligned} \tag{5.1.2}$$

where we have used $(1+s)p + q = 1 + sp$, since $p + q = 1$. Note that (5.1.6) is the same as (5.1.4) with $p' = p_{i+1}, p = p_i$, and $f_A g_A / f_a g_a = 1 + s$. Fixed points of (5.1.6) are determined from $p' = p$. We find two fixed points: $p_* = 0$, corresponding to a population in which allele A is absent; and $p_* = 1$, corresponding to a population in which allele A is fixed. Intuitively, $p_* = 0$ is unstable while $p_* = 1$ is stable.

To illustrate how a stability analysis is performed analytically for a difference equation (instead of a differential equation), consider the general difference equation

$$p' = f(p)$$

With $p = p_*$ a fixed point such that $p_* = f(p_*)$, we write $p = p_* + \epsilon$ so that (5.1.7) becomes

$$\begin{aligned} p_* + \epsilon' &= f(p_* + \epsilon) \\ &= f(p_*) + \epsilon f'(p_*) + \dots \\ &= p_* + \epsilon f'(p_*) + \dots , \end{aligned}$$

where $f'(p_*)$ denotes the derivative of f evaluated at p_*. Therefore, to leading-order in ϵ

$$|\epsilon'/\epsilon| = |f'(p_*)| ,$$

and the fixed point is stable provided that $|f'(p_*)| < 1$. For our haploid model,

$$f(p) = \frac{(1+s)p}{1+sp}, \quad f'(p) = \frac{1+s}{(1+sp)^2}$$

Table 5.3: A haploid genetic model of mutation-selection balance.

genotype	A	a
freq. of gamete	p	q
relative fitness	1	$1-s$
freq after selection	p/w	$(1-s)q/w$
freq after mutation	$(1-u)p/w$	$[(1-s)q + up]/w$
normalization	$w = p + (1-s)q$	

so that $f'(p_* = 0) = 1 + s > 1$, and $f'(p_* = 1) = 1/(1+s) < 1$, confirming that $p_* = 0$ is unstable and $p_* = 1$ is stable.

If the selection coefficient s is small, the model equation (5.1.6) simplifies further. We have

$$\begin{aligned} p' &= \frac{(1+s)p}{1+sp} \\ &= (1+s)p\left(1 - sp + \mathrm{O}\left(s^2\right)\right) \\ &= p + \left(p - p^2\right)s + \mathrm{O}\left(s^2\right) \end{aligned}$$

so that to leading-order in s,

$$p' - p = sp(1-p)$$

If $p' - p \ll 1$, which is valid for $s \ll 1$, we can approximate this difference equation by the differential equation

$$dp/dn = sp(1-p)$$

which shows that the frequency of allele A satisfies the now very familiar logistic equation.

Although a polymorphism for this gene exists in the population as the new allele spreads, eventually A becomes fixed in the population and the polymorphism is lost. In the next section, we consider how a polymorphism can be maintained in a haploid population by a balance between mutation and selection.

Mutation-selection balance

We consider a gene with two alleles: a wildtype allele A and a mutant allele a. We view the mutant allele as a defective genotype, which confers on the carrier a lowered fitness $1-s$ relative to the wildtype. Although all mutant alleles may not have identical DNA sequences, we assume that they share in common the same phenotype of reduced fitness. We model the opposing effects of two evolutionary forces: natural selection, which favors the wildtype allele A over the mutant allele a, and mutation, which confers a small probability u that allele A mutates to allele a in each newborn individual. Schematically,

$$A \underset{s}{\overset{u}{\rightleftharpoons}} a$$

where u represents mutation and s represents selection. The model is shown in Table 5.3. The equations for p and q in the next generation are

$$\begin{aligned} p' &= \frac{(1-u)p}{w} \\ &= \frac{(1-u)p}{1-s(1-p)} \end{aligned}$$

Table 5.4: The terminology of diploidy.

genotype	AA	Aa	aa
referred to as	wildtype homozygote	heterozygote	mutant homozygote
frequency	P	Q	R

and

$$\begin{aligned} q' &= \frac{(1-s)q + up}{w} \\ &= \frac{(1-s-u)q + u}{1-sq} \end{aligned} \tag{5.1.3}$$

where we have used $p+q=1$ to eliminate q from the equation for p' and p from the equation for q'. The equations for p' and q' are linearly dependent since $p'+q'=1$, and we need solve only one of them.

Considering (5.1.13), the fixed points determined from $p'=p$ are $p_*=0$, for which the mutant allele a is fixed in the population and there is no polymorphism, and the solution to

$$1 - s\,(1-p_*) = 1-u$$

which is $p_* = 1-u/s$, and there is a polymorphism. The stabilities of these two fixed points are determined by considering $p' = f(p)$, with $f(p)$ given by the righthand-side of (5.1.13). Taking the derivative of f,

$$f'(p) = \frac{(1-u)(1-s)}{[1-s(1-p)]^2}$$

so that

$$f'\,(p_*=0) = \frac{1-u}{1-s}, \quad f'\,(p_*=1-u/s) = \frac{1-s}{1-u}.$$

Applying the criterion $|f'(p_*)| < 1$ for stability, $p_* = 0$ is stable for $s < u$ and $p_* = 1 - u/s$ is stable for $s > u$. A polymorphism is therefore possible under mutation-selection balance when $s > u > 0$.

This page titled 5.1: Haploid Genetics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

5.2: Diploid Genetics

Most sexually reproducing species are diploid. In particular, our species Homo sapiens is diploid with two exceptions: we are haploid at the gamete stage (sperm and unfertilized egg); and males are haploid for most genes on the unmatched X and Y sex chromosomes (females are XX and diploid). This latter seemingly innocent fact is of great significance to males suffering from genetic diseases due to an X-linked recessive mutation inherited from their mother. Females inheriting this mutation are most probably disease-free because of the functional gene inherited from their father.

A polymorphic gene with alleles A and a can appear in a diploid gene as three distinct genotypes: AA, Aa and aa. Conventionally, we denote A to be the wildtype allele and a the mutant allele. Table 5.4 presents the terminology of diploidy.

As for haploid genetics, we will determine evolution equations for allele and/or genotype frequencies. To develop the appropriate definitions and relations, we initially assume a population of size N (which we will later take to be infinite), and assume that the number of individuals with genotypes AA, Aa and aa are N_{AA}, N_{Aa} and N_{aa}. Now, $N = N_{AA} + N_{Aa} + N_{aa}$. Define genotype frequencies P, Q and R as

$$P = \frac{N_{AA}}{N}, \quad Q = \frac{N_{Aa}}{N}, \quad R = \frac{N_{aa}}{N}$$

so that $P + Q + R = 1$. It will also be useful to define allele frequencies. Let n_A and n_a be the number of alleles A and a in the population, with $n = n_A + n_a$ the total number of alleles. Since the population is of size N and diploidy, $n = 2N$; and since each homozygote contains two identical alleles, and each heterozygote contains one of each allele, $n_A = 2N_{AA} + N_{Aa}$ and $n_a = 2N_{aa} + N_{Aa}$. Defining the allele frequencies p and q as previously,

$$\begin{aligned} p &= n_A/n \\ &= \frac{2N_{AA} + N_{Aa}}{2N} \\ &= P + \frac{1}{2}Q \end{aligned}$$

and similarly,

$$\begin{aligned} q &= n_a/n \\ &= \frac{2N_{aa} + N_{Aa}}{2N} \\ &= R + \frac{1}{2}Q. \end{aligned}$$

With five frequencies, P, Q, R, p, q, and four constraints $P + Q + R = 1, p + q = 1, p = P + Q/2, q = R + Q/2$, how many independent frequencies are there? In fact, there are two because one of the four constraints is linearly dependent. We may choose any two frequencies other than the choice $\{p, q\}$ as our linearly independent set. For instance, one choice is $\{P, p\}$; then,

$$q = 1 - p, \quad Q = 2(p - P), \quad R = 1 + P - 2p$$

Similarly, another choice is $\{P, Q\}$; then

$$R = 1 - P - Q, \quad p = P + \frac{1}{2}Q, \quad q = 1 - P - \frac{1}{2}Q$$

Sexual reproduction

Diploid reproduction may be sexual or asexual, and sexual reproduction may be of varying types (e.g., random mating, selfing, brother-sister mating, and various other types of assortative mating). The two simplest types to model exactly are random mating and selfing. These mating systems are useful for contrasting the biology of both outbreeding and inbreeding.

Random mating

Random mating is perhaps the simplest mating system to model. Here, we assume a well-mixed population of individuals that have equal probability of mating with every other individual. We will determine the genotype frequencies of the zygotes (fertilized eggs) in terms of the allele frequencies using two approaches: (1) the gene pool approach, and (2) the mating table approach.

Table 5.5: Random mating table.

mating	frequency	AA	Aa	aa
$AA \times AA$	P^2	P^2	0	0
$AA \times Aa$	$2PQ$	PQ	PQ	0
$AA \times aa$	$2PR$	0	$2PR$	0
$Aa \times Aa$	Q^2	$\frac{1}{4}Q^2$	$\frac{1}{2}Q^2$	$\frac{1}{4}Q^2$
$Aa \times aa$	$2QR$	0	QR	QR
$aa \times aa$	R^2	0	0	R^2
Totals	$(P+Q+R)^2$	$\left(P+\frac{1}{2}Q\right)^2$	$2\left(P+\frac{1}{2}Q\right)\left(R+\frac{1}{2}Q\right)$	$\left(R+\frac{1}{2}Q\right)^2$
	$=1$	$=p^2$	$=2pq$	$=q^2$

The gene pool approach models sexual reproduction by assuming that males and females release their gametes into pools. Offspring genotypes are determined by randomly combining one gamete from the male pool and one gamete from the female pool. As the probability of a random gamete containing allele A or a is equal to the allele's population frequency p or q, respectively, the probability of an offspring being AA is p^2, of being Aa is $2pq$ (male A female $a+$ female A male a), and of being aa is q^2. Therefore, after a single generation of random mating, the genotype frequencies can be given in terms of the allele frequencies by

$$P = p^2, \quad Q = 2pq, \quad R = q^2$$

This is the celebrated Hardy-Weinberg law. Notice that under the assumption of random mating, there is now only a single independent frequency, greatly simplifying the mathematical modeling. For example, if p is taken as the independent frequency, then

$$q = 1 - p, \quad P = p^2, \quad Q = 2p(1-p), \quad R = (1-p)^2$$

Most modeling is done assuming random mating unless the biology under study is influenced by inbreeding.

The second approach uses a mating table (see Table 5.5). This approach to modeling sexual reproduction is more general and can be applied to other mating systems. We explain this approach by considering the mating $AA \times Aa$. The genotypes AA and Aa have frequencies P and Q, respectively. The frequency of AA males mating with Aa females is PQ and is the same as AA females mating with Aa males, so the sum is $2PQ$. Half of the offspring will be AA and half Aa, and the frequencies PQ are denoted under progeny frequency. The sums of all the progeny frequencies are given in the Totals row, and the random mating results are recovered upon use of the relationship between the genotype and allele frequencies.

Selfing

Perhaps the next simplest type of mating system is self-fertilization, or selfing. Here, an individual reproduces sexually (passing through a haploid gamete stage in its life-cycle), but provides both of the gametes. For example, the nematode worm $C.$ elegans can reproduce by selfing. The mating table for selfing is given in Table 5.6. The selfing frequency of a particular genotype is just the frequency of the genotype itself. For a selfing population, disregarding selection or any other evolutionary

Table 5.6: Selfing mating table.

mating	frequency	AA	Aa	aa
$AA\otimes$	P	P	0	0
$Aa\otimes$	Q	$\frac{1}{4}Q$	$\frac{1}{2}Q$	$\frac{1}{4}Q$
$aa\otimes$	R	0	0	R
Totals	1	$P+\frac{1}{4}Q$	$\frac{1}{2}Q$	$R+\frac{1}{4}Q$

forces, the genotype frequencies evolve as

$$P' = P + \frac{1}{4}Q, \quad Q' = \frac{1}{2}Q, \quad R' = R + \frac{1}{4}Q.$$

Assuming an initially heterozygous population, we solve (5.2.6) with the initial conditions $Q_0 = 1$ and $P_0 = R_0 = 0$. In the worm lab, this type of initial population is commonly created by crossing wildtype homozygous C. elegans males with mutant homozygous C. elegans hermaphrodites, where the mutant allele is recessive. Wildtype hermaphrodite offspring, which are necessarily heterozygous, are then picked to separate worm plates and allowed to self-fertilize. (Do you see why the experiment is not done with wildtype hermaphrodites and mutant males?) From the equation for Q' in $(5.2.6)$, we have $Q_n = (1/2)^n$, and from symmetry, $P_n = R_n$. Then, since $P_n + Q_n + R_n = 1$, we obtain the complete solution

$$P_n = \frac{1}{2}\left(1 - \left(\frac{1}{2}\right)^n\right), \quad Q_n = \left(\frac{1}{2}\right)^n, \quad R_n = \frac{1}{2}\left(1 - \left(\frac{1}{2}\right)^n\right)$$

The main result to be emphasized here is that the heterozygosity of the population decreases by a factor of two in each generation. Selfing populations rapidly become homozygous.

Constancy of allele frequencies

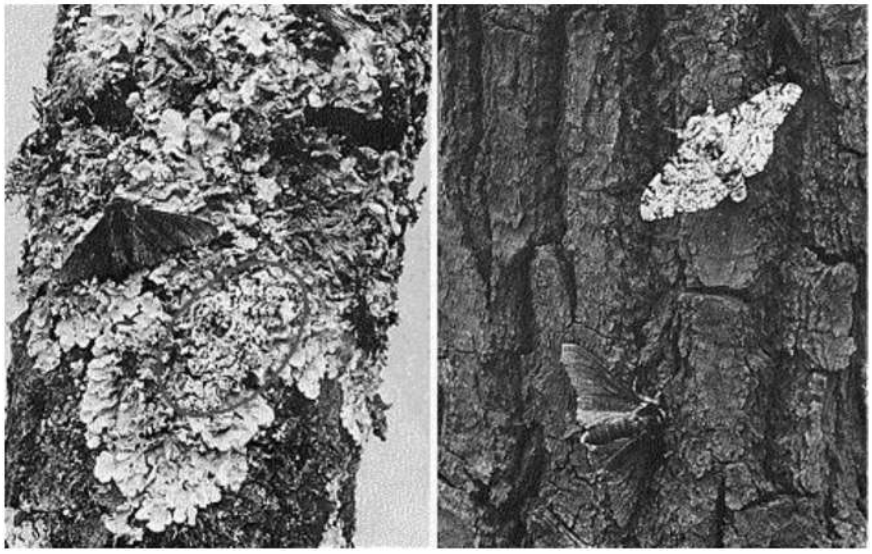


Figure 5.1: Evolution of peppered moths in industrializing England (from H. B. D. Kettlewell).

These results are not general, however, and it is possible to construct other mating systems for which allele frequencies do change.

Nevertheless, the conservation of allele frequencies by random mating is an important element of neo-Darwinism. In Darwin's time, most biologists believed in blending inheritance, where the genetic material from parents with different traits actually blended in their offspring, rather like the mixing of paints of different colors. If blending inheritance occurred, then genetic variation, or polymorphism, would eventually be lost over several generations as the "genetic paints" became well-mixed. Mendel's work on peas, published in 1866, suggested a particulate theory of inheritance, where the genetic material, later called genes, maintain their integrity across generations. Sadly, Mendel's paper was not read by Darwin (who published The Origin of Species in 1859 and died in 1882) or other influential biologists during Mendel's lifetime (Mendel died in 1884). After being rediscovered in 1900, Mendel and his work eventually became widely celebrated.

Spread of a favored allele

We consider the spread of a favored allele in a diploid population. The classic example - widely repeated in biology textbooks as a modern example of natural selection - is the change in the frequencies of the dark and light phenotypes of the peppered moth during England's industrial revolution. The evolutionary story begins with the observation that pollution killed the light colored lichen on trees during industrialization of the cities. On the one hand, light colored peppered moths camouflage well on light colored lichens, but are exposed to birds on plain tree bark. On the other hand, dark colored peppered moths camouflage well on plain tree bark, but are exposed on light colored lichens (see Fig. 5.1). Natural selection therefore favored the light-colored allele in preindustrialized England and the dark-colored allele during industrialization. It is believed that the dark-colored allele increased rapidly under natural selection in industrializing England.

We present our model in Table 5.7. Here, we consider aa as the wildtype genotype and normalize its fitness to unity. The allele A is the mutant whose frequency increases in the population. In our example of the peppered moth, the aa pheno-

Table 5.7: A diploid genetic model of the spread of a favored allele assuming random mating.

genotype	AA	Aa	aa
freq. of zygote	p^2	$2pq$	q^2
relative fitness	$1+s$	$1+sh$	1
freq after selection	$(1+s)p^2/w$	$2(1+sh)pq/w$	q^2/w
normalization	$w=(1+s)p^2+2(1+sh)pq+q^2$		

type is light colored and the AA phenotype is dark colored. The color of the Aa phenotype depends on the relative dominance of A and a. Usually, no pigment results in light color and is a consequence of nonfunctioning pigment-producing genes. One functioning pigment-producing allele is usually sufficient to result in a dark-colored moth. With A a functioning pigment-producing allele and a the mutated nonfunctioning allele, a is most likely recessive, A is most likely dominant, and the phenotype of Aa is most likely dark, so $h \approx 1$. For the moment, though, we leave h as a free parameter.

We assume random mating, and this simplification is used to write the genotype frequencies as $P=p^2, Q=2pq$, and $R=q^2$. Since $q=1-p$, we reduce our problem to determining an equation for p' in terms of p. Using $p'=P_s+(1/2)Q_s$, where p' is the A allele frequency in the next generation's zygotes, and P_S and Q_S are the AA and Aa genotype frequencies, respectively, in the present generation after selection,

$$p' = \frac{(1+s)p^2+(1+sh)pq}{w}$$

where $q=1-p$, and

$$\begin{aligned} w &= (1+s)p^2+2(1+sh)pq+q^2 \\ &= 1+s\left(p^2+2hpq\right) \end{aligned}$$

After some algebra, the final evolution equation written solely in terms of p is

$$p' = \frac{(1+sh)p+s(1-h)p^2}{1+2shp+s(1-2h)p^2}$$

The expected fixed points of this equation are $p_*=0$ (unstable) and $p_*=1$ (stable), where our assignment of stability assumes positive selection coefficients.

The evolution equation (5.2.9) in this form is not particularly illuminating. In general, a numerical solution would require specifying numerical values for s and h, as well as an initial value for p. Here, to determine how the spread of A depends on the dominance coefficient h, we investigate analytically the increase of A assuming $s \ll 1$. We Taylor-series expand the right-hand-side of (5.2.9) in powers of s, keeping terms to order s :

$$\begin{aligned} p' &= \frac{(1+sh)p+s(1-h)p^2}{1+2shp+s(1-2h)p^2} \\ &= \frac{p+s\left(hp+(1-h)p^2\right)}{1+s\left(2hp+(1-2h)p^2\right)} \\ &= \left(p+s\left(hp+(1-h)p^2\right)\right)\left(1-s\left(2hp+(1-2h)p^2\right)+\mathrm{O}\left(s^2\right)\right) \\ &= p+sp\left(h+(1-3h)p-(1-2h)p^2\right)+\mathrm{O}\left(s^2\right) \end{aligned} \tag{5.2.1}$$

Table 5.8: Seven common monogenic diseases.

disease	mutation	symptoms
Thalassemia	haemoglobin	anemia
Sickle cell anemia	haemoglobin	anemia
Haemophilia	blood clotting factor	uncontrolled bleeding
Cystic Fibrosis	chloride ion channel	thick lung mucous

disease	mutation	symptoms
Tay-Sachs disease	Hexosaminidase A enzyme	nerve cell damage
Fragile X syndrome	FMR1 gene	mental retardation
Huntington's disease	HD gene	brain degeneration

If $s \ll 1$, we expect a small change in allele frequency in each generation, so we can approximate $p' - p \approx dp/dn$, where n denotes the generation number, and $p = p(n)$. The approximate differential equation obtained from (5.2.10) is

$$\frac{dp}{dn} = sp\left(h + (1-3h)p - (1-2h)p^2\right)$$

If A is partially dominant so that $h \neq 0 (e.g.$, the heterozygous moth is darker than the homozgygous mutant moth), then the solution to (5.2.11) behaves similarly to the solution of a logistic equation: p initially grows exponentially as $p(n) = p_0 \exp(shn)$, and asymptotes to one for large n. If A is recessive so that $h = 0 (e.g.$,, the heterozygous moth is as light-colored as the homozygous mutant moth), then (5.2.11) reduces to

$$\frac{dp}{dn} = sp^2(1-p), \quad \text{for } h = 0$$

Of main interest is the initial growth of p when $p(0) = p_0 \ll 1$, so that $dp/dn \approx sp^2$. This differential equation may be integrated by separating variables to yield

$$\begin{aligned} p(n) &= \frac{p_0}{1 - sp_0 n} \\ &\approx p_0\left(1 + sp_0 n\right) \end{aligned}$$

The frequency of a recessive favored allele increases only linearly across generations, a consequence of the heterozygote being hidden from natural selection. Most likely, the peppered-moth heterozygote is significantly darker than the light-colored homozygote since the dark colored moth rapidly increased in frequency over a short period of time.

As a final comment, linear growth in the frequency of A when $h = 0$ is sensitive to our assumption of random mating. If selfing occurred, or another type of close family mating, then a recessive favored allele may still increase exponentially. In this circumstance, the production of homozygous offspring from more frequent heterozygote pairings allows selection to act more effectively.

Mutation-selection balance

By virtue of self-knowledge, the species with the most known mutant phenotypes is Homo sapiens. There are thousands of known genetic diseases in humans, many of them caused by mutation of a single gene (called a monogenic disease). For an easy-to-read overview of genetic disease in humans, see the website

http://www.who.int/genomics/public/geneticdiseases

Table 5.9: A diploid genetic model of mutation-selection balance assuming random mating.

genotype	AA	Aa	aa
freq. of zygote	p^2	$2pq$	q^2
relative fitness	1	$1-sh$	$1-s$
freq after selection	p^2/w	$2(1-sh)pq/w$	$(1-s)q^2/w$
normalization	$w = p^2 + 2(1-sh)pq + (1-s)q^2$		

Table 5.8 lists seven common monogenic diseases. The first two diseases are maintained at significant frequencies in some human populations by heterosis. We will discuss in $5.2.4 the maintenance of a polymorphism by heterosis, for which the heterozygote has higher fitness than either homozygote. It is postulated that Tay-Sachs disease, prevalent among ancestors of Eastern European Jews, and cystic fibrosis may also have been maintained by heterosis acting in the past. (Note that the cystic fibrosis gene was

identified in 1989 by a Toronto group led by Lap Chee Tsui, who later became President of the University of Hong Kong.) The other disease genes listed may be maintained by mutation-selection balance.

Our model for diploid mutation-selection balance is given in Table 5.9. We further assume that mutations of type $A \to$ a occur in gamete production with frequency u. Back-mutation is neglected. The gametic frequency of A and a after selection but before mutation is given by $\hat{p} = P_s + Q_s/2$ and $\hat{q} = R_s + Q_s/2$, and the gametic frequency of a after mutation is given by $q' = u\hat{p} + \hat{q}$. Therefore,

$$q' = \left(u\left(p^2 + (1-sh)pq\right) + \left((1-s)q^2 + (1-sh)pq\right)\right)/w$$

where

$$\begin{aligned} w &= p^2 + 2(1-sh)pq + (1-s)q^2 \\ &= 1 - sq(2hp + q) \end{aligned}$$

Using $p = 1 - q$, we write the evolution equation for q' in terms of q alone. After some algebra that could be facilitated using a computer algebra software such as Mathematica, we obtain

$$q' = \frac{u + (1 - u - \mathrm{sh}(1+u))q - s(1 - h(1+u))q^2}{1 - 2\,\mathrm{sh}\,q - s(1-2h)q^2}$$

To determine the equilibrium solutions of (5.2.14), we set $q_* \equiv q' = q$ to obtain a cubic equation for q_*. Because of the neglect of back mutation in our model, one solution readily found is $q_* = 1$, in which all the A alleles have mutated to a. The $q_* = 1$ solution may be factored out of the cubic equation resulting in a quadratic equation, with two solutions. Rather than show the exact result here, we determine equilibrium solutions under two approximations: (i) $0 < u \ll h, s$, and; (ii) $0 = h < u < s$.

First, when $0 < u \ll h, s$, we look for a solution of the form $q_* = au + \mathrm{O}\left(u^2\right)$, with a constant, and Taylor series expand in u (assuming $s, h = \mathrm{O}\left(u^0\right)$). If such a solution exists, then (5.2.14) will determine the unknown coefficient a. We have

$$\begin{aligned} au + \mathrm{O}\left(u^2\right) &= \frac{u + (1 - \mathrm{sh})au + \mathrm{O}\left(u^2\right)}{1 - 2\mathrm{shau} + \mathrm{O}\left(u^2\right)} \\ &= (1 + a - \mathrm{sh}\,a)u + \mathrm{O}\left(u^2\right) \end{aligned}$$

Table 5.10: Equilibrium frequencies of the genotypes at the diploid mutationselection balance.

genotype	AA	Aa	aa
freq: $0 < u \ll s, h$	$1 + \mathrm{O}(u)$	$2u/sh + \mathrm{O}\left(u^2\right)$	$u^2/(sh)^2 + \mathrm{O}\left(u^3\right)$
freq: $0 = h < u < s$	$1 + \mathrm{O}(\sqrt{u})$	$2\sqrt{u/s} + \mathrm{O}(u)$	u/s

and equating powers of u, we find $a = 1 + a - \mathrm{sh}\,a$, or $a = 1/\mathrm{sh}$. Therefore,

$$q_* = u/sh + \mathrm{O}\left(u^2\right), \qquad \text{for } 0 < u \ll h, s.$$

Second, when $0 = h < u < s$, we substitute $h = 0$ directly into (5.2.14),

$$q_* = \frac{u + (1-u)q_* - sq_*^2}{1 - sq_*^2}$$

which we then write as a cubic equation q_*

$$q_*^3 - q_*^2 - \frac{u}{s}q_* + \frac{u}{s} = 0$$

By factoring this cubic equation, we find

$$(q_* - 1)\left(q_*^2 - u/s\right) = 0$$

and the polymorphic equilibrium solution is

$$q_* = \sqrt{u/s}, \qquad \text{for } 0 = h < u < s$$

Because $q_* < 1$ only if $s > u$, this solution does not exist if $s < u$.

Table 5.10 summarizes our results for the equilibrium frequencies of the genotypes at mutation-selection balance. The first row of frequencies, $0 < u \ll s, h$, corresponds to a dominant $(h = 1)$ or partially-dominant $(u \ll h < 1)$ mutation, where the heterozygote is of reduced fitness and shows symptoms of the genetic disease. The second row of frequencies, $0 = h < u < s$, corresponds to a recessive mutation, where the heterozygote is symptom-free. Notice that individuals carrying a dominant mutation are twice as prevalent in the population as individuals homozygous for a recessive mutation (with the same u and s).

A heterozygote carrying a dominant mutation most commonly arises either de novo (by direct mutation of allele A) or by the mating of a heterozygote with a wildtype. The latter is more common for $s \ll 1$, while the former must occur for $s = 1$ (a heterozygote with an $s = h = 1$ mutation by definition does not reproduce). One of the most common autosomal dominant genetic diseases is Huntington's disease, resulting in brain deterioration during middle age. Because individuals with Huntington's disease have children before disease symptoms appear, s is small and the disease is usually passed to offspring by the mating of a (heterozygote) with a wildtype homozygote. For a recessive mutation, a mutant homozygote usually occurs by the mating of two heterozygotes. If both parents carry a single recessive disease allele, then their child has a $1/4$ chance of getting the disease.

Heterosis

Heterosis, also called overdominance or heterozygote advantage, occurs when the heterozygote has higher fitness than either homozygote. The best-known examples

Table 5.11: A diploid genetic model of heterosis assuming random mating.

genotype	AA	Aa	aa
freq. of zygote	p^2	$2pq$	q^2
relative fitness	$1-s$	1	$1-t$
freq after selection	$(1-s)p^2/w$	$2pq/w$	$(1-t)q^2/w$
normalization	$w = (1-s)p^2 + 2pq + (1-t)q^2$		

are sickle-cell anemia and thalassemia, diseases that both affect hemoglobin, the oxygen-carrier protein of red blood cells. The sickle-cell mutations are most common in people of West African descent, while the thalassemia mutations are most common in people from the Mediterranean and Asia. In Hong Kong, the television stations occasionally play public service announcements concerning thalassemia. The heterozygote carrier of the sickle-cell or thalassemia gene is healthy and resistant to malaria; the wildtype homozygote is healthy, but susceptible to malaria; the mutant homozygote is sick with anemia. In class, we will watch the short video, A Mutation Story, about the sickle cell gene.

Table 5.11 presents our model of heterosis. Both homozygotes are of lower fitness than the heterozygote, whose relative fitness we arbitrarily set to unity. Writing the equation for p', we have

$$\begin{aligned} p' &= \frac{(1-s)p^2 + pq}{1 - sp^2 - tq^2} \\ &= \frac{p - sp^2}{1 - t + 2tp - (s+t)p^2} \end{aligned}$$

At equilibrium, $p_* \equiv p' = p$, and we obtain a cubic equation for p_* :

$$(s+t)p_*^3 - (s+2t)p_*^2 + tp_* = 0$$

Evidently, $p_* = 0$ and $p_* = 1$ are fixed points, and (5.2.20) can be factored as

$$p(1-p)(t - (s+t)p) = 0$$

The polymorphic solution is therefore

$$p_* = \frac{t}{s+t}, \quad q_* = \frac{s}{s+t},$$

valid when $s, t > 0$. Since the value of q_* can be large, recessive mutations that cause disease, yet are highly prevalent in a population, are suspected to provide some benefit to the heterozygote. However, only a few genes are unequivocally known to exhibit heterosis.

This page titled 5.2: Diploid Genetics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

5.3: Frequency-Dependent Selection

A polymorphism may also result from frequency-dependent selection. A wellknown model of frequency-dependent selection is the Hawk-Dove game. Most commonly, frequency-dependent selection is studied using game theory, and following John Maynard Smith, one looks for an evolutionarily stable strategy (ESS).

We consider two phenotypes: Hawk and Dove, with no mating between different phenotypes (for example, different phenotypes may correspond to different

player \ opponent	H	D
H	$E_{HH} = -2$	$E_{HD} = 2$
D	$E_{DH} = 0$	$E_{DD} = 1$

Table 5.12: General payoff matrix for the Hawk-Dove game, and the usually assumed values. The payoffs are payed to the player (first column) when playing against the opponent (first row).

species, such as hawks and doves). We describe the Hawk-Dove game as follows: (i) when Hawk meets Dove, Hawk gets the resource and Dove retreats before injury; (ii) when two Hawks meet, they engage in an escalating fight, seriously risking injury, and; (iii) when two Doves meet, they share the resource.

The Hawk-Dove game is modeled by a payoff matrix, as shown in Table 5.12. The player in the first column receives the payoff when playing the opponent in the first row. For instance, Hawk playing Dove gets the payoff E_{HD}. The numerical values are commonly chosen such that $E_{HH} < E_{DH} < E_{DD} < E_{HD}$, that is, Hawk playing Dove does better than Dove playing Dove does better than Dove playing Hawk does better than Hawk playing Hawk.

Frequency-dependent selection occurs because the expected payoff to a Hawk or a Dove depends on the frequency of Hawks and Doves in the population. For example, a Hawk in a population of Doves does well, but a Hawk in a population of Hawks does poorly.

A population of all Doves is unstable to invasion by Hawks (because Hawk playing against Dove does better than Dove playing against Dove), and similarly a population of all Hawks is unstable to invasion by Doves. These two possible equilibria are therefore unstable, and the stable equilibrium consists of a mixed population of Hawks and Doves. In game theory, this mixed equilibrium is called a mixed Nash equilibrium, and is determined by assuming that the expected payoff to a Hawk in a mixed population of Hawks and Doves is the same as the expected payoff to a Dove.

With p the frequency of Hawks and q the frequency of Doves, the expected payoff to a Hawk is $pE_{HH} + qE_{HD}$, and the expected payoff to a Dove is $pE_{DH} + qE_{DD}$, so that the mixed Nash equilibrium satisfies

$$pE_{HH} + qE_{HD} = pE_{DH} + qE_{DD}$$

Substituting in $q = 1 - p$ and solving for p, we obtain

$$p = \frac{E_{HD} - E_{DD}}{(E_{HD} - E_{DD}) + (E_{DH} - E_{HH})}$$

and with the numerical values in Table 5.12,

$$\begin{aligned} p_* &= \frac{2-1}{(2-1)+(0+2)} \\ &= 1/3. \end{aligned}$$

Thus the stable polymorphic population maintained by frequency-dependent selection consists of $1/3$ Hawks and $2/3$ Doves.

This page titled 5.3: Frequency-Dependent Selection is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

5.4: Linkage Equilibrium

When considering a polymorphism at a single genetic locus, we assumed two distinct alleles, A and a. The diploid then occurs as one of three types: AA, Aa and aa. We now consider a polymorphism at two genetic loci, each with two distinct alleles. If the alleles at the first genetic loci are A and a, and those at the second B and b, then four distinct haploid gametes are possible, namely AB, Ab, aB and ab. Ten distinct diplotypes are possible, obtained by forming pairs of all possible haplotypes. We can write these ten diplotypes as $AB/AB, AB/Ab, AB/aB, AB/ab$, $Ab/Ab, Ab/aB, Ab/ab, aB/aB, aB/ab$, and ab/ab, where the numerator represents the haplotype from one parent, the denominator represents the haplotype from the other parent. We do not distinguish here which haplotype came from which parent.

To proceed further, we define the allelic and gametic frequencies for our two loci problem in Table 5.13. If the probability that a gamete contains allele A or a does not depend on whether the gamete contains allele B or b, then the two loci are said to be independent. Under the assumption of independence, the gametic frequencies are the products of the allelic frequencies, i.e., $p_{AB} = p_A p_B, p_{Ab} = p_A p_b$, etc.

Often, the two loci are not independent. This can be due to epistatic selection, or epistasis. As an example, suppose that two loci in humans influence height, and that the most fit genotype is the one resulting in an average height. Selection that favors the average population value of a trait is called normalizing or stabilizing. Suppose that A and B are hypothetical tall alleles, a and b are short alleles, and a person with two tall and two short alleles obtains average height. Then selection may favor the specific genotypes $AB/ab, Ab/Ab, Ab/aB$, and aB/aB. Selection may act against both the genotypes yielding above average heights, $AB/AB, AB/Ab$, and AB/aB, and those yielding below average heights, $Ab/ab, aB/ab$ and ab/ab. Epistatic selection occurs because the fitness of the A, a loci depends on which alleles are present at the B, b loci. Here, A has higher fitness when paired with b than when paired with B.

The two loci may also not be independent because of a finite population size (i.e., stochastic effects). For instance, suppose a mutation $a \to A$ occurs only once in a finite population (in an infinite population, any possible mutation occurs an infinite number of times), and that A is strongly favored by natural selection. The frequency of A may then increase. If a nearby polymorphic locus on the same chromosome as A happens to be B (say, with a polymorphism b in the population), then AB gametes may substantially increase in frequency, with Ab absent. We say that the allele B hitchhikes with the favored allele A.

When the two loci are not independent, we say that the loci are in gametic phase disequilibrium, or more commonly linkage disequilibrium, sometimes abbreviated as LD. When the loci are independent, we say they are in linkage equilibrium. Here, we will model how two loci, initially in linkage disequilibrium, approach linkage equilibrium through the process of recombination.

To begin, we need a rudimentary understanding of meiosis. During meiosis, a

Table 5.13: Definitions of allelic and gametic frequencies for two genetic loci each with two alleles.

allele or gamete genotype	A	a	B	b	AB	Ab	aB	ab
frequency	p_A	p_a	p_B	p_b	p_{AB}	p_{Ab}	p_{aB}	p_{ab}

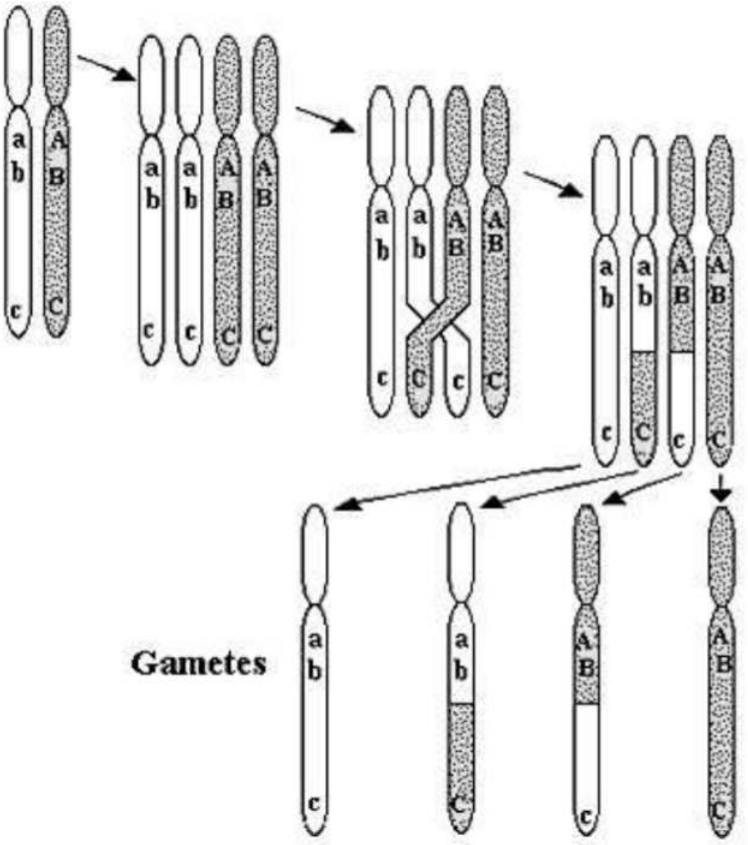


Figure 5.2: A schematic of crossing-over and recombination during meiosis (figure from Access Excellence @ the National Health Museum)

diploid cell's DNA, arranged in very long molecules called chromosomes, is replicated once and separated twice, producing four haploid cells, each containing half of the original cell's chromosomes. Sexual reproduction results in syngamy, the fusing of a haploid egg and sperm cell to form a diploid zygote cell.

Fig. 5.2 presents a schematic of meiosis and the process of crossing-over resulting in recombination. In a diploid, each chromosome has a corresponding sister chromosome, one chromosome originating from the egg, one from the sperm. These sibling chromosomes have the same genes, but possibly different alleles. In Fig. 5.2, we schematically show the alleles a, b, c on the light chromosome, and the alleles A, B, C on its sister's dark chromosome. In the first step of meiosis, each chromosome replicates itself exactly. In the second step, sister chromosomes exchange genetic material by the process of crossing-over. All four chromosomes then separate into haploid cells. Notice from the schematic that the process of crossing-over can result in genetic recombination. Suppose that the schematic of Fig. 5.2 represents the production of sperm by a male. If the chromosome from the male's father contains the alleles ABC and that from the male's mother abc, recombination can result in the sperm containing a chromosome with alleles ABc (the third gamete in Fig. 5.2). We say this chromosome is a recombinant; it contains alleles from both its paternal grandfather and paternal grandmother. It is likely that the precise combination of alleles on this recombinant chromosome has never existed before in a single person. Recombination is the reason why everybody, with the exception of identical twins, is genetically unique.

Genes that occur on the same chromosome are said to be linked. The closer the genes are to each other on the chromosome, the tighter the linkage, and the less likely recombination will separate them. Tightly linked genes are likely to be inherited from the same grandparent. Genes on different chromosomes are by definition unlinked; independent assortment of chromosomes results in a 50% chance of a gamete receiving either grandparents' genes. To define and model the evolution of linkage disequilibrium, we first obtain allele frequencies from gametic frequencies by

$$p_A = p_{AB} + p_{Ab}, \quad p_a = p_{aB} + p_{ab}$$
$$p_B = p_{AB} + p_{aB}, \quad p_b = p_{Ab} + p_{ab}$$

Since the frequencies sum to unity,

$$p_A + p_a = 1, \quad p_B + p_b = 1, \quad p_{AB} + p_{Ab} + p_{aB} + p_{ab} = 1.$$

There are three independent gametic frequencies and only two independent allelic frequencies, so in general it is not possible to obtain the gametic frequencies from the allelic frequencies without assuming an additional constraint such as linkage equilibrium. We can, however, introduce an additional variable D, called the coefficient of linkage disequilibrium, and define D to be the difference between the gametic frequency p_{AB} and what this gametic frequency would be if the loci were in linkage equilibrium:

$$p_{AB} = p_A p_B + D$$

Using $p_{AB} + p_{Ab} = p_A$ to eliminate p_{AB} in (5.4.3), we obtain

$$p_{Ab} = p_A p_b - D$$

Likewise, using $p_{AB} + p_{aB} = p_B$

$$p_{aB} = p_a p_B - D$$

and using $p_{ab} + p_{ab} = p_a$

$$p_{ab} = p_a p_b + D.$$

With our definition, positive linkage disequilibrium $(D > 0)$ implies excessive AB and ab gametes and deficient Ab and aB gametes; negative linkage disequilibrium $(D < 0)$ implies the opposite. D attains its maximum value of $1/4$ when $p_{AB} = p_{ab} = 1/2$, and attains its minimum value of $-1/4$ when $p_{Ab} = p_{aB} = 1/2$. An equality obtainable from (5.4.3, 5.4.4, 5.4.5, 5.4.6) that we will later find useful is

$$\begin{aligned} p_{AB}p_{ab} - p_{Ab}p_{aB} &= (p_A p_B + D)(p_a p_b + D) - (p_A p_b - D)(p_a p_B - D) \\ &= D(p_A p_B + p_a p_b + p_A p_b + p_a p_B) \\ &= D \end{aligned} \tag{5.4.1}$$

Without selection and mutation, D evolves only because of recombination. With primes representing the values in the next generation, and using $p'_A = p_A$ and $p'_B = p_B$ because sexual reproduction by itself does not change allele frequencies,

$$\begin{aligned} D' &= p'_{AB} - p'_A p'_B \\ &= p'_{AB} - p_A p_B \\ &= p'_{AB} - (p_{AB} - D) \\ &= D + (p'_{AB} - p_{AB}), \end{aligned}$$

where we have used (5.4.3) to obtain the third equality. The change in D is therefore equal to the change in frequency of the AB gametes,

$$D' - D = p'_{AB} - p_{AB}$$

Table 5.14: Computation of gamete frequencies.

		gamete freq / diploid freq			
diploid	**dip freq**	**AB**	**Ab**	**aB**	**ab**
AB/AB	p_{AB}^2	1	0	0	0
AB/Ab	$2p_{AB}p_{Ab}$	1/2	1/2	0	0
AB/aB	$2p_{AB}p_{aB}$	1/2	0	1/2	0
AB/ab	$2p_{AB}p_{ab}$	$(1-r)/2$	$r/2$	$r/2$	$(1-r)/2$
Ab/Ab	p_{Ab}^2	0	1	0	0
Ab/aB	$2p_{Ab}p_{aB}$	$r/2$	$(1-r)/2$	$(1-r)/2$	$r/2$
Ab/ab	$2p_{Ab}p_{ab}$	0	1/2	0	1/2
aB/aB	p_{aB}^2	0	0	1	0

aB/ab	$2p_{aB}p_{ab}$	0	0	$1/2$	$1/2$
ab/ab	p_{ab}^2	0	0	0	1

To understand why gametic frequencies change across generations, we should first recognize when they do not change. Without genetic recombination, chromosomes maintain their exact identity across generations. Chromosome frequencies without recombination are therefore constant, and for genetic loci on the same chromosome with alleles A, a and B, b, say, $p'_{AB} = p_{AB}$. In an infinite population without selection or mutation, gametic frequencies change only for genetic loci in linkage disequilibrium on different chromosomes, or for genetic loci in linkage disequilibrium on the same chromosome subjected to genetic recombination.

We will compute the frequency p'_{AB} of AB gametes in the next generation, given the frequency p_{AB} of AB gametes in the present generation, using two different methods. The first method uses a mating table. The second method makes a direct probability argument.

The mating table is shown in Table 5.14. The first column is the parent diplotype before meiosis. The second column is the diplotype frequency assuming random mating. The next four columns are the haploid genotype frequencies (normalized by the corresponding diploid frequencies to simplify the table presentation). Here, we define r to be the frequency at which the gamete arises from a combination of grandmother and grandfather genes. If the A, a and B, b loci occur on the same chromosome, then r is the recombination frequency due to crossing-over. If the A, a and B, b loci occur on different chromosomes, then because of the independent assortment of chromosomes there is an equal probability that the gamete contains all grandfather or grandmother genes, or contains a combination of grandmother and grandfather genes, so that $r = 1/2$. Notice that crossing-over or independent assortment is of importance for those pairs of genes for which the grandfather's and grandmother's contribution to the diploid genotype share no common alleles (i.e., AB/ab and Ab/aB genotypes). The frequency p'_{AB} in the next generation is given by the sum of the AB column (after multiplication by the diploid frequencies). Therefore,

$$\begin{aligned} p'_{AB} &= p_{AB}^2 + p_{AB}p_{Ab} + p_{AB}p_{aB} + (1-r)p_{AB}p_{ab} + rp_{Ab}p_{aB} \\ &= p_{AB}\left(p_{AB} + p_{Ab} + p_{aB} + p_{ab}\right) + r\left(p_{Ab}p_{aB} - p_{AB}p_{ab}\right) \\ &= p_{AB} - rD \end{aligned} \tag{5.4.2}$$

where the final equality makes use of (5.4.2) and (5.4.7).

The second method for computing p'_{AB} is more direct. An AB haplotype can arise from a diploid of general type AB/XX without recombination, or a diploid of type AX/XB with recombination. Therefore,

$$p'_{AB} = (1-r)p_{AB} + rp_Ap_B$$

where the first term is from non-recombinants and the second term from recombinants. With $p_Ap_B = p_{AB} - D$, we have

$$\begin{aligned} p'_{AB} &= (1-r)p_{AB} + r\left(p_{AB} - D\right) \\ &= p_{AB} - rD \end{aligned}$$

the same result as (5.4.9).

Using (5.4.8) and (5.4.9), we derive

$$D' = (1-r)D,$$

with the solution

$$D_n = D_0(1-r)^n$$

Recombination decreases linkage disequilibrium in each generation by a factor of $(1-r)$. Tightly linked genes on the same chromosome have small values of r; unlinked genes on different chromosomes have $r = 1/2$. For unlinked genes, linkage disequilibrium decreases by a factor of two in each generation. We conclude that very strong selection is required to maintain linkage disequilibrium for genes on different chromosomes, while weak selection can maintain linkage disequilibrium for tightly linked genes.

This page titled 5.4: Linkage Equilibrium is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

5.5: Random Genetic Drift

Up to now, our simplified genetic models have all assumed an infinite population, neglecting stochastic effects. Here, we consider a finite population: the resulting stochastic effects on the allelic frequencies is called random genetic drift. The simplest genetic model incorporating random genetic drift assumes a fixed-sized population of N individuals, and models the evolution of a diallelic haploid genetic locus.

There are two widely used genetic models for finite populations: the WrightFisher model and the Moran model. The Wright-Fisher model is most similar to our infinite-population discrete-generation model. In the Wright-Fisher model, N adult individuals release a very large number of gametes into a gene pool, and the next generation is formed from N random gametes independently chosen from the gene pool. The Moran model takes a different approach. In the Moran model, a single evolution step consists of one random individual in the population reproducing, and another random individual dying, with the population always maintained at the constant size N. Because two random events occur every generation in the Moran model, and N random events occur every generation in the Wright-Fisher model, a total of $N/2$ evolution steps in the Moran model is comparable, but not exactly identical, to a single discrete generation in the Wright-Fisher model. It has been shown, however, that the two models become identical in the limit of large N. For our purposes, the Moran model is mathematically more tractable and we adopt it here.

We develop our model analogously to the stochastic population growth model derived in §3.1. We let n denote the number of A-alleles in the population, and $N-n$ the number of a-alleles. With $n=0,1,2,\ldots,N$ a discrete random variable, the probability mass function $p_n(g)$ denotes the probability of having nA-alleles at evolution step g. With n A-alleles in a population of size N, the probability of an individual with A either reproducing or dying is given by $s_n=n/N$; the corresponding probability for a is $1-s_n$. There are three ways to obtain a population of nA-alleles at evolution step $g+1$. First, there were nA-alleles at evolution step g, and the individual reproducing carried the same allele as the individual dying. Second, there were $n-1$-alleles at evolution step g, and the individual reproducing has A and the individual dying has a. And third, there were $n+1A$-alleles at evolution step g, and the individual reproducing has a and the individual dying has A. Multiplying the probabilities and summing the three cases results in

$$p_n(g+1)=\left(s_n^2+(1-s_n)^2\right)p_n(g)+s_{n-1}\left(1-s_{n-1}\right)p_{n-1}(g) + s_{n+1}\left(1-s_{n+1}\right)p_{n+1}(g) \tag{5.5.1}$$

Note that this equation is valid for $0<n<N$, and that the equations at the boundaries-representing the probabilities that one of the alleles is fixed-are

$$p_0(g+1)=p_0(g)+s_1\left(1-s_1\right)p_1(g) \tag{5.5.2}$$

$$p_N(g+1)=p_N(g)+s_{N-1}\left(1-s_{N-1}\right)p_{N-1}(g) \tag{5.5.3}$$

The boundaries are called absorbing and the probability of fixation of an allele monotonically increases with each birth and death. Once the probability of fixation of an allele is unity, there are no further changes in allele frequencies.

We illustrate the solution of (5.5.1)-(5.5.3) in Fig. 5.3 for a small population of size $N=20$, and where the number of A-alleles is precisely known in the founding generation, with either (a) $p_{10}(0)=1$, or; (b) $p_{13}(0)=1$. We plot the probability mass density of the number of A-alleles every N evolution steps, up to $7N$ steps, corresponding to approximately fourteen discrete generations of evolution in the Wright-Fisher model. Notice how the probability distribution diffuses away from its initial value, and how the probabilities eventually concentrate on the boundaries, with both p_0 and p_{20} monotonically increasing. In fact, after a large number of generations, p_0 approaches the initial frequency of the a-allele and p_{20} approaches the initial frequency of the A-allele (not shown in figures).

Figure 5.3: p_n versus n with $N = 20$. Evolution steps plotted correspond to $g = 0, N, 2N, \ldots, 7N$. (a) Initial number of A individuals is $n = 10$. (b) Initial number of A individuals is $n = 13$. and note that $S(x) = x$. Similarly, we have $s_{n+1} = S(x + \Delta x)$ and $s_{n-1} = S(x - \Delta x)$, where $\Delta x = 1/N$. Then, with $\Delta t = 2/N$, (5.5.4) transforms into<\p>

$$P(x, t + \Delta t) - P(x, t) = S(x + \Delta x)(1 - S(x + \Delta x))P(x + \Delta x, t) - 2S(x)(1 - S(x))P(x, t) + S(x - \Delta x)(1 - S(x - \Delta x))P(x - \Delta x, t) \quad (5.5.4)$$

To simplify further, we use the well-known central-difference approximation to the second-derivative of a function $f(x)$,

$$f''(x) = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{\Delta x^2} + \mathrm{O}\left(\Delta x^2\right)$$

and recognize the right-hand side of (5.5.7) to be the numerator of a second-derivative. With $\Delta x = \Delta t/2 = 1/N \to 0$, and $S(x) = x$, we derive to leading-order in $1/N$ the partial differential equation

$$\frac{\partial P(x, t)}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial x^2}(V(x)P(x, t))$$

with

$$V(x) = \frac{x(1 - x)}{N}$$

To interpret the meaning of the function $V(x)$ we will use the following result from probability theory. For n independent trials, each with probability of success p and probability of failure $1 - p$, the number of successes, denoted by X, is a binomial random variable with parameters (n, p). Well-known results are $E[X] = np$ and $\mathrm{Var}[X] = np(1 - p)$, where $E[...]$ is the expected value, and $\mathrm{Var}[...]$ is the variance.

Now, the number of A-alleles chosen when forming the next Wright-Fisher generation is a binomial random variable n' with parameters $(N, n/N)$. Therefore, $E[n'] = n$ and $\mathrm{Var}[n'] = n(1 - n/N)$. With $x' = n'/N$ and $x = n/N$, we have $E[x'] = x$, and $\mathrm{Var}[x'] = x(1 - x)/N$. The function $V(x)$ can therefore be interpreted as the variance of x over a single Wright-Fisher generation.

Although (5.5.9) and (5.5.10) depend explicitly on the population size N, the population size can be eliminated by a simple change of variables. If we let

$$\tau = t/N,$$

then the differential equation (5.5.9) transforms to

$$\frac{\partial P(x, \tau)}{\partial \tau} = \frac{1}{2}\frac{\partial^2}{\partial x^2}(x(1 - x)P(x, \tau))$$

independent of N. The change-of-variables given by (5.5.11) simply states that a doubling of the population size will lengthen the time scale of evolution by a corresponding factor of two. Remember that here we are already working under the assumption that population sizes are large.

Equation (5.5.12) is a diffusion-like equation for the probability distribution function $P(x, \tau)$. A diffusion approximation for studying genetic drift was first introduced into population genetics by its founders Fisher and Wright, and was later extensively developed in the $1950'$ s by the Japanese biologist Motoo Kimura. Here, our analysis of this equation relies on a more recent paper by McKane and Waxman (2007), who showed how to construct the analytical solution of (5.5.12) at the boundaries of x. For $0 < x < 1$, it is suggestive from the numerical solutions shown in Fig. 5.3 that the probability distribution might asymptotically become independent of x. Accordingly, we look for an asymptotic solution to (5.5.12) at the interior values of x satisfying $P(x, \tau) = P(\tau)$. Equation (5.5.12) becomes

$$\frac{dP(\tau)}{d\tau} = \frac{1}{2}(x(1 - x))''P(\tau) = -P(\tau)$$

with asymptotic solution

$$P(\tau) = ce^{-\tau}$$

and we observe that the total probability over the interior values of x decays exponentially.

To understand how the solution behaves at the boundaries of x, we require boundary conditions for $P(x,\tau)$. In fact, because $P(x,\tau)$ is singular at the boundaries of x, appropriate boundary conditions can not be obtained directly from the difference equations given by (5.5.2 and 5.5.3). Rather, boundary conditions are most easily obtained by first recasting (5.5.12) into the form of a continuity equation.

We let $j(x,\tau)$ denote the so-called probability current. In a small region of size Δx lying in the interval $(x, x+\Delta x)$, the time-rate of change of the probability $P(x,\tau)\Delta x$ is due to the flow of probability into and out of this region. With an appropriate definition of the probability current, we have in general

$$\frac{\partial}{\partial\tau}(P(x,\tau)\Delta x) = j(x,t) - j(x+\Delta x)$$

or as $\Delta x \to 0$,

$$\frac{\partial P(x,\tau)}{\partial\tau} + \frac{\partial j(x,\tau)}{\partial x} = 0$$

which is the usual form of a continuity equation. Identification of (5.5.15) with (5.5.12) shows that the probability current of our problem is given by

$$j(x,\tau) = -\frac{1}{2}\frac{\partial}{\partial x}(x(1-x)P(x,\tau)).$$

Now, since the total probability is unity; that is,

$$\int_0^1 P(x,\tau)dx = 1$$

probability can not flow in or out of the boundaries of x, and we must therefore have

$$j(0,\tau) = 0, \quad j(1,\tau) = 0,$$

which are the appropriate boundary conditions for (5.5.12).

We can look for stationary solutions of (5.5.12). Use of the continuity equation (5.5.15) together with the boundary conditions (5.5.18) shows that the stationary solution has zero probability current; that is, $j(x) = 0$. Integration of $j(x)$ using (5.5.16) results in

$$x(1-x)P(x) = c_1$$

where c_1 is an integration constant. The readily apparent solution is $P(x) = c_1/[x(1-x)]$, but there are also two less obvious solutions. Probability distribution functions are allowed to contain singular solutions corresponding to Dirac delta functions, and here we consider the possibility of there being Dirac delta functions at the boundaries. By noticing that both $x\delta(x) = 0$ and $(1-x)\delta(1-x) = 0$, we can see that the general solution for $P(x)$ can be written as

$$P(x) = \frac{c_1}{x(1-x)} + c_2\delta(x) + c_3\delta(1-x).$$

Requiring $P(x)$ to satisfy (5.5.17) results in $c_1 = 0$ and $c_2 + c_3 = 1$. To determine the remaining free constant, we can compute the mean frequency of the A-allele in the population. From the continuity equation $(5.5.15)$, we have

$$\begin{aligned}\frac{\partial}{\partial\tau}\int_0^1 xP(x,\tau)dx &= -\int_0^1 x\frac{\partial j(x,\tau)}{\partial x}dx \\ &= \int_0^1 j(x,\tau)dx \\ &= 0\end{aligned}$$

where the first integral on the right-hand-side was done by parts using the boundary conditions given by (5.5.18), and the second integral was done using (5.5.16) and the vanishing of $x(1-x)P(x)$ on the boundaries. The mean frequency of the A allele is therefore a constant-as one would expect for a nondirectional random genetic drift-and we can assume that its initial value is p. We therefore obtain for our stationary solution

$$P(x) = (1-p)\delta(x) + p\delta(1-x)$$

The eventual probability of fixing the A-allele is therefore simply equal to its initial frequency. For example, suppose that within a population of N individuals homogeneous for a, a single neutral mutation occurs so that one individual now carries the A-allele. What is the probability that the A-allele eventually becomes fixed? Our result would yield the probability $1/N$, which is the initial frequency of the A-allele. Intuitively, after a sufficient number of generations has passed, all living individuals should be descendant from a single ancestral individual living at the time the single mutation occurred. The probability that that single individual carried the A-allele is just $1/N$.

We further note that Kimura was the first to find an analytical solution of the diffusion equation (5.5.12). A solution method using Fourier transforms can be found in the appendix of McKane and Waxman (2007). In addition to making use of Dirac delta functions, these authors require Heaviside step functions, Bessel functions, spherical Bessel functions, hypergeometric functions, Legendre polynomials, and Gegenbauer polynomials. The resulting solution is of the form

$$P(x,\tau) = \Pi_0(\tau)\delta(x) + \Pi_1(\tau)\delta(1-x) + f(x,\tau)$$

If the number of A-alleles are known at the initial instant, and the frequency is p, then

$$P(x,0) = \delta(x-p)$$

and $\Pi_0(0) = \Pi_1(0) = 0; f(x,0) = \delta(x-p)$. As $\tau \to \infty$, we have $f(x,\tau) \to 0$; $\Pi_0(\tau) \to 1-p$ and $\Pi_1(\tau) \to p$

We can easily demonstrate at least one exact solution of the form (5.5.22). For an initial uniform probability distribution given by

$$P(x,0) = 1,$$

the probability distribution at the interior points remains uniform and is given by $(5.5.13)$ with $c = 1$. If we assume the form of solution given by (5.5.22) together with the requirement of unit probability given by (5.5.17), we can obtain the exact result

$$P(x,\tau) = \frac{1}{2}\left(1-e^{-\tau}\right)\left(\delta(x) + \delta(1-x)\right) + e^{-\tau}$$

This page titled 5.5: Random Genetic Drift is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

6: Biochemical Reactions

Biochemistry is the study of the chemistry of life. It can be considered a branch of molecular biology, perhaps more focused on specific molecules and their reactions, or a branch of chemistry focused on the complex chemical reactions occurring in living organisms. One can guess that the first application of biochemistry happened about 5000 years ago when bread was made using yeast.

Modern biochemistry, however, had a relatively slow start among the sciences, as did modern biology. Isaac Newton's publication of Principia Mathematica in 1687 preceded Darwin's Origin of Species in 1859 by almost 200 years. I find this amazing because the ideas of Darwin are in many ways simpler and easier to understand than the mathematical theory of Newton. Most of the delay must be attributed to a fundamental conflict between science and religion. The physical sciences experienced this conflict early-witness the famous prosecution of Galileo by the Catholic Church in 1633, during which Galileo was forced to recant his heliocentric viewbut the conflict of religion with evolutionary biology continues even to this day. Advances in biochemistry were initially delayed because it was long believed that life was not subject to the laws of science the way non-life was, and that only living things could produce the molecules of life. Certainly, this was more a religious conviction than a scientific one. Then Friedrich Wöhler in 1828 published his landmark paper on the synthesis of urea (a waste product neutralizing toxic ammonia before excretion in the urine), demonstrating for the first time that organic compounds can be created artificially.

Here, we present mathematical models for some important biochemical reactions. We begin by introducing a useful model for a chemical reaction: the law of mass action. We then model what may be the most important biochemical reactions, namely those catalyzed by enzymes. Using the mathematical model of enzyme kinetics, we consider three fundamental enzymatic properties: competitive inhibition, allosteric inhibition, and cooperativity.

6.1: The Law of Mass Action
6.2: Enzyme Kinetics
6.3: Competitive Inhibition
6.4: Allosteric Inhibition
6.5: Cooperativity

This page titled 6: Biochemical Reactions is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

6.1: The Law of Mass Action

The law of mass action describes the rate at which chemicals interact in reactions. It is assumed that different chemical molecules come into contact by collision before reacting, and that the collision rate is directly proportional to the number of molecules of each reacting species. Suppose that two chemicals A and B react to form a product chemical C, written as

$$A + B \xrightarrow{k} C,$$

with k the rate constant of the reaction. For simplicity, we will use the same symbol C, say, to refer to both the chemical C and its concentration. The law of mass action says that dC/dt is proportional to the product of the concentrations A and B, with proportionality constant k. That is,

$$\frac{dC}{dt} = kAB \tag{6.1.1}$$

Similarly, the law of mass action enables us to write equations for the time-derivatives of the reactant concentrations A and B :

$$\frac{dA}{dt} = -kAB, \quad \frac{dB}{dt} = -kAB$$

Notice that when using the law of mass action to find the rate-of-change of a concentration, the chemical that the arrow points towards is increasing in concentration (positive sign), the chemical that the arrow points away from is decreasing in concentration (negative sign). The product of concentrations on the right-hand-side is always that of the reactants from which the arrow points away, multiplied by the rate constant that is on top of the arrow.

Equation 6.1.1 can be solved analytically using conservation laws. Each reactant, original and converted to product, is conserved since one molecule of each reactant gets converted into one molecule of product. Therefore,

$$\begin{aligned} &\frac{d}{dt}(A + C) = 0 \qquad && A + C = A_0 \\ &\frac{d}{dt}(B + C) = 0 \qquad \Longrightarrow && B + C = B_0 \end{aligned}$$

where A_0 and B_0 are the initial concentrations of the reactants, and no product is present initially. Using the conservation laws, Equation 6.1.1 becomes

$$\frac{dC}{dt} = k\left(A_0 - C\right)\left(B_0 - C\right), \text{ with } C(0) = 0$$

which may be integrated by separating variables. After some algebra, the solution is determined to be

$$C(t) = A_0 B_0 \frac{e^{(B_0 - A_0)kt} - 1}{B_0 e^{(B_0 - A_0)kt} - A_0}$$

which is a complicated expression with the simple limits

$$\lim_{t\to\infty} C(t) = \begin{cases} A_0 & \text{if } A_0 < B_0 \\ B_0 & \text{if } B_0 < A_0 \end{cases}$$

The reaction stops after one of the reactants is depleted; and the final concentration of the product is equal to the initial concentration of the depleted reactant.

If we also include the reverse reaction,

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C_1$$

then the time-derivative of the product is given by

$$\frac{dC}{dt} = k_+ AB - k_- C$$

Notice that $k_+$and k_-have different units. At equilibrium, $\dot{C} = 0$, and using the conservation laws $A + C = A_0, B + C = B_0$, we obtain

$$(A_0 - C)(B_0 - C) - \frac{k_-}{k_+} C = 0$$

from which we define the equilibrium constant K_{eq} by

$$K_{eq} = k_-/k_+$$

which has units of concentration. Therefore, at equilibrium, the concentration of the product is given by the solution of the quadratic equation

$$C^2 - (A_0 + B_0 + K_{eq})\, C + A_0 B_0 = 0$$

with the extra condition that $0 < C < \min(A_0, B_0)$. For instance, if $A_0 = B_0 \equiv R_0$, then at equilibrium,

$$C = R_0 - \frac{1}{2} K_{eq} \left(\sqrt{1 + 4R_0/K_{eq}} - 1 \right)$$

If $K_{eq} \ll R_0$, then A and B have a high affinity, and the reaction proceeds mainly to C, with $C \to R_0$.

Below are two interesting reactions. In reaction (ii), A is assumed to be held at a constant concentration.

$$A + X \underset{k_-}{\overset{k_+}{\rightleftarrows}} 2X$$

(ii)

$$A + X \overset{k_1}{\to} 2X, \quad X + Y \overset{k_2}{\to} 2Y, \quad Y \overset{k_3}{\to} B$$

Can you write down the equations for $\dot{X}$ in reaction (i), and $\dot{X}$ and $\dot{Y}$ in reaction (ii)? When normalized properly, the equations from reaction (ii) reduce to the LotkaVolterra predator-prey equations introduced in §1.4. The chemical concentrations X and Y, therefore, oscillate in time like predators and their prey.

This page titled 6.1: The Law of Mass Action is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

6.2: Enzyme Kinetics

Enzymes are catalysts, usually proteins, that help convert other molecules called substrates into products, but are themselves unchanged by the reaction. Each enzyme has high specificity for at least one reaction, and it can accelerate this reaction by millions of times. Without enzymes, most biochemical reactions are too slow for life to be possible. Enzymes are so important to our lives that a single amino acid mutation in one enzyme out of the more than 2000 enzymes in our bodies can result in a severe or lethal genetic disease.

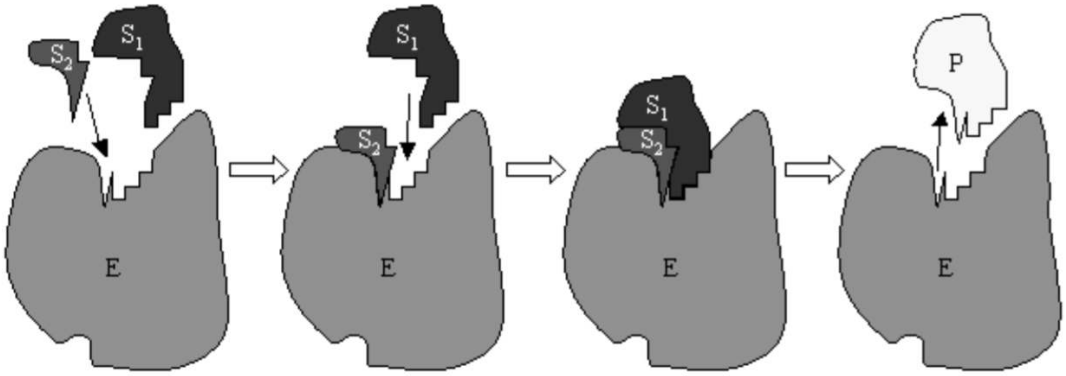


Figure 6.1: A Michaelis-Menten reaction of two substrates converting to one product. (Drawn by User:IMeowbot, released under the GNU Free Documentation License.)

where C is a complex formed by the enzyme and the substrate. A cartoon of the Michaelis-Menten reaction with an enzyme catalyzing a reaction between two substrates is shown in Fig. 6.1. Commonly, substrate is continuously provided to the reaction and product is continuously removed. The removal of product has been modeled by neglecting the reverse reaction $P + E \to C$. A continuous provision of substrate allows us to assume that S is held at an approximately constant concentration.

The differential equations for C and P can be obtained from the law of mass action:

$$dC/dt = k_1 SE - (k_{-1} + k_2)\, C$$
$$dP/dt = k_2 C.$$

Biochemists usually want to determine the reaction velocity dP/dt in terms of the substrate concentration S and the total enzyme concentration E_0. We can eliminate E in favor of E_0 from the conservation law that the enzyme, free and bound, is conserved; that is

$$\frac{d(E+C)}{dt} = 0 \quad \Longrightarrow \quad E + C = E_0 \quad \Longrightarrow \quad E = E_0 - C$$

and we can rewrite the equation for dC/dt eliminating E :

$$\begin{aligned} \frac{dC}{dt} &= k_1 S\,(E_0 - C) - (k_{-1} + k_2)\, C \\ &= k_1 E_0 S - (k_{-1} + k_2 + k_1 S)\, C \end{aligned} \tag{6.2.1}$$

Because S is assumed to be held constant, the complex C is expected to be in equilibrium, with the rate of formation equal to the rate of dissociation. With this so-called quasi-steady-state approximation, we may assume that $C = 0$ in $(6.3.4)$, and we have

$$C = \frac{k_1 E_0 S}{k_{-1} + k_2 + k_1 S}$$

The reaction velocity is then given by

$$\begin{aligned} \frac{dP}{dt} &= k_2 C \\ &= \frac{k_1 k_2 E_0 S}{k_{-1} + k_2 + k_1 S} \\ &= \frac{V_m S}{K_m + S} \end{aligned} \tag{6.2.2}$$

where two fundamental constants are defined:

$$K_m = (k_{-1} + k_2)/k_1, \quad V_m = k_2 E_0.$$

The Michaelis-Menten constant or the Michaelis constant K_m has units of concentration, and the maximum reaction velocity V_m has units of concentration divided by time. The interpretation of these constants is obtained by considering the following limits:

$$\begin{aligned} &\text{as } S \to \infty, \quad C \to E_0 \text{ and } dP/dt \to V_m \\ &\text{if } S = K_m, \quad C = \tfrac{1}{2} E_0 \text{ and } dP/dt = \tfrac{1}{2} V_m. \end{aligned}$$

Therefore, V_m is the limiting reaction velocity obtained by saturating the reaction with substrate so that every enzyme is bound; and K_m is the concentration of S at which only one-half of the enzymes are bound and the reaction proceeds at one-half maximum velocity.

This page titled 6.2: Enzyme Kinetics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

6.3: Competitive Inhibition

Competitive inhibition occurs when inhibitor molecules compete with substrate molecules for binding to the same enzyme's active site. When an inhibitor is bound to the enzyme, no product is produced so competitive inhibition will reduce the velocity of the reaction. A cartoon of this process is shown in Fig. 6.2.

To model competitive inhibition, we introduce an additional reaction associated with the inhibitor-enzyme binding:

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C_1 \xrightarrow{k_2} P + E,$$

$$I + E \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} C_2.$$

With more complicated enzymatic reactions, the reaction schematic becomes difficult to interpret. Perhaps an easier way to visualize the reaction is from the following redrawn schematic:

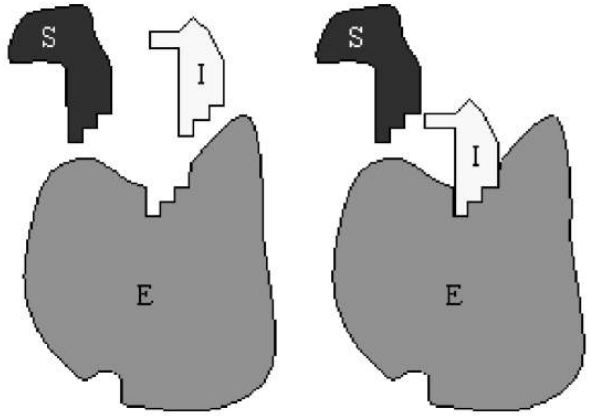


Figure 6.2: Competitive inhibition. (Drawn by G. Andruk, released under the GNU Free Documentation License.)

$$E \underset{k_{-1}}{\overset{k_1 S}{\rightleftarrows}} C_1 \xrightarrow{k_2} P + E$$

$$k_3 I \downarrow \quad \uparrow k_{-3}$$

$$C_2$$

Here, the substrate S and inhibitor I are combined with the relevant rate constants, rather than treated separately. It is immediately obvious from this redrawn schematic that inhibition is accomplished by sequestering enzyme in the form of C_2 and preventing its participation in the catalysis of S to P.

Our goal is to determine the reaction velocity $\dot{P}$ in terms of the substrate and inhibitor concentrations, and the total concentration of the enzyme (free and bound). The law of mass action applied to the two complexes and the product results in

$$\frac{dC_1}{dt} = k_1 SE - (k_{-1} + k_2) C_1$$
$$\frac{dC_2}{dt} = k_3 IE - k_{-3} C_2$$
$$\frac{dP}{dt} = k_2 C_1$$

The enzyme, free and bound, is conserved so that

$$\frac{d}{dt}(E + C_1 + C_2) = 0 \quad \Longrightarrow \quad E + C_1 + C_2 = E_0 \quad \Longrightarrow \quad E = E_0 - C_1 - C_2.$$

Under the quasi-equilibrium approximation, $C_1 = \dot{C}_2 = 0$, so that

$$k_1 S(E_0 - C_1 - C_2) - (k_{-1} + k_2) C_1 = 0,$$
$$k_3 I(E_0 - C_1 - C_2) - k_{-3} C_2 = 0,$$

which results in the following system of two linear equations and two unknowns (C_1 and C_2) :

$$(k_{-1} + k_2 + k_1 S) C_1 + k_1 S C_2 = k_1 E_0 S \tag{6.3.1}$$
$$k_3 I C_1 + (k_{-3} + k_3 I) C_2 = k_3 E_0 I \tag{6.3.2}$$

We define the Michaelis-Menten constant K_m as before, and an additional constant K_i associated with the inhibitor reaction:

$$K_m = \frac{k_{-1} + k_2}{k_1}, \quad K_i = \frac{k_{-3}}{k_3}$$

Dividing (6.3.2) by k_1 and (6.3.3) by k_3 yields

$$(K_m + S) C_1 + S C_2 = E_0 S \tag{6.3.3}$$
$$I C_1 + (K_i + I) C_2 = E_0 I \tag{6.3.4}$$

Since our goal is to obtain the velocity of the reaction, which requires determining C_1, we multiply (6.3.5) by $(K_i + I)$ and (6.3.6) by S, and subtract:

$$(K_m + S)(K_i + I) C_1 + S(K_i + I) C_2 = E_0 (K_i + I) S$$
$$S I C_1 + S(K_i + I) C_2 = E_0 S I$$
$$((K_m + S)(K_i + I) - S I) C_1 = K_i E_0 S$$

or after cancellation and rearrangement

$$C_1 = \frac{K_i E_0 S}{K_m K_i + K_i S + K_m I}$$
$$= \frac{E_0 S}{K_m (1 + I/K_i) + S}$$

Therefore, the reaction velocity is given by

$$\frac{dP}{dt} = \frac{(k_2 E_0) S}{K_m (1 + I/K_i) + S}$$
$$= \frac{V_m S}{K'_m + S} \tag{6.3.5}$$

where

$$V_m = k_2 E_0, \quad K'_m = K_m (1 + I/K_i).$$

By comparing the inhibited reaction velocity (6.3.7) and (6.3.8) with the uninhibited reaction velocity (6.2.6) and (6.2.7), we observe that inhibition increases the Michaelis-Menten constant of the reaction, but leaves unchanged the maximum reaction velocity. Since the Michaelis-Menten constant is defined as the substrate concentration required to attain one-half of the maximum reaction velocity, addition of an inhibitor with a fixed substrate concentration acts to decrease the reaction velocity. However, a reaction saturated with substrate still attains the uninhibited maximum reaction velocity.

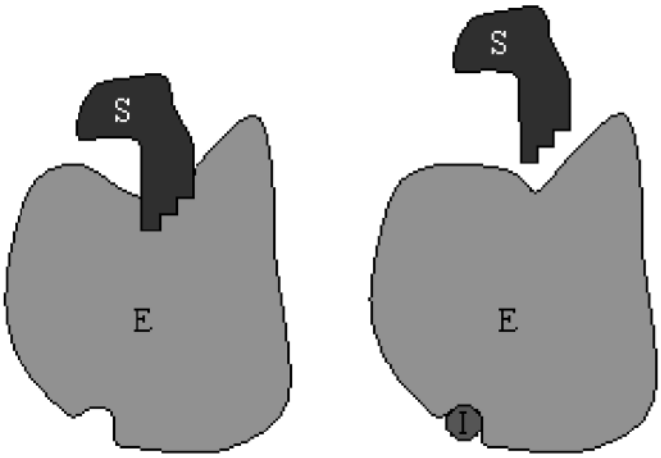


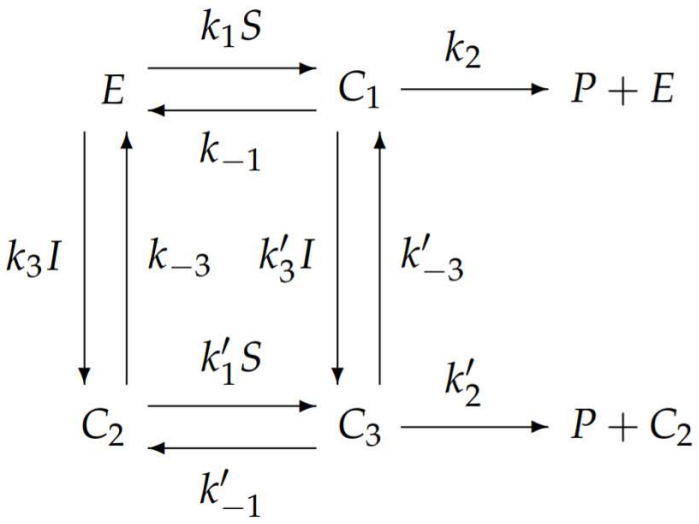
Figure 6.3: Allosteric inhibition. (Unknown artist, released under the GNU Free Documentation License.)

This page titled 6.3: Competitive Inhibition is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

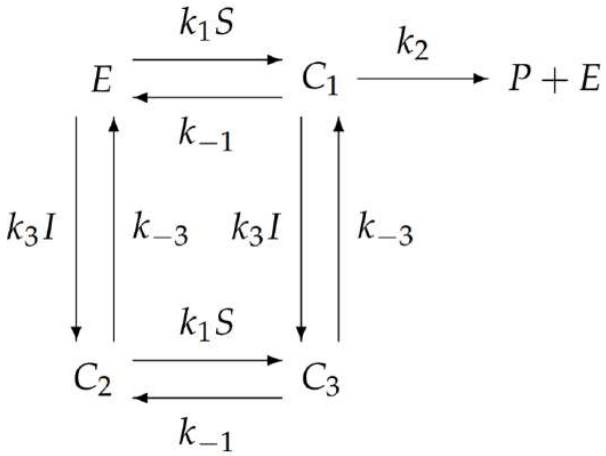
6.4: Allosteric Inhibition

The term allostery comes from the Greek word allos, meaning different, and stereos, meaning solid, and refers to an enzyme with a regulatory binding site separate from its active binding site. In our model of allosteric inhibition, an inhibitor molecule is assumed to bind to its own regulatory site on the enzyme, resulting in either a lowered binding affinity of the substrate to the enzyme, or a lowered conversion rate of substrate to product. A cartoon of allosteric inhibition due to a lowered binding affinity is shown in Fig. 6.3.

In general, we need to define three complexes: C_1 is the complex formed from substrate and enzyme; C_2 from inhibitor and enzyme, and; C_3 from substrate, inhibitor, and enzyme. We write the chemical reactions as follows:



The general model for allosteric inhibition with ten independent rate constants appears too complicated to analyze. We will simplify this general model to one with fewer rate constants that still exhibits the unique features of allosteric inhibition. One possible but uninteresting simplification assumes that if I binds to E, then S does not; however, this reduces allosteric inhibition to competitive inhibition and loses the essence of allostery. Instead, we simplify by allowing both I and S to simultaneously bind to E, but we assume that the binding of I prevents substrate conversion to product. With this simplification, $k'_2 = 0$. To further reduce the number of independent rate constants, we assume that the binding of S to E is unaffected by the bound presence of I, and the binding of I to E is unaffected by the bound presence of S. These approximations imply that all the primed rate constants equal the corresponding unprimed rate constants, e.g., $k'_1 = k_1$, etc. With these simplifications, the schematic of the chemical reaction simplifies to



and now there are only five independent rate constants. We write the equations for the complexes using the law of mass action:

$$dC_1$$

$$\frac{dC_1}{dt} = k_1 SE + k_{-3} C_3 - (k_{-1} + k_2 + k_3 I)\, C_1 \tag{6.4.1}$$

$$\frac{dC_2}{dt} = k_3 IE + k_{-1} C_3 - (k_{-3} + k_1 S)\, C_2 \tag{6.4.2}$$

$$\frac{dC_3}{dt} = k_3 IC_1 + k_1 SC_2 - (k_{-1} + k_{-3})\, C_3 \tag{6.4.3}$$

while the reaction velocity is given by

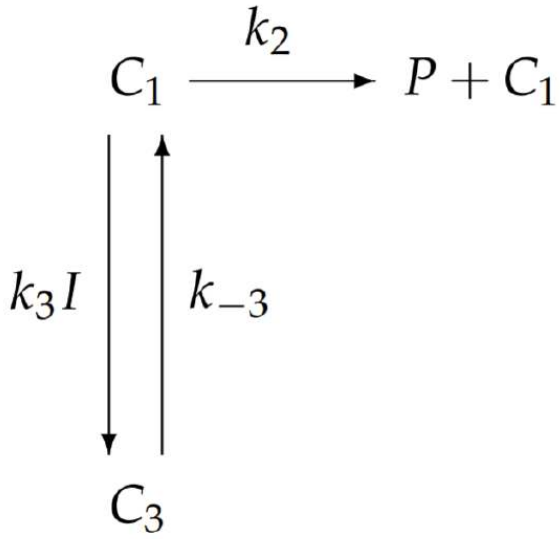
$$\frac{dP}{dt} = k_2 C_1$$

Again, both free and bound enzyme is conserved, so that $E = E_0 - C_1 - C_2 - C_3$. With the quasi-equilibrium approximation $\dot{C}_1 = \dot{C}_2 = \dot{C}_3 = 0$, we obtain a system of three equations and three unknowns: C_1, C_2 and C_3. Despite our simplifications, the analytical solution for the reaction velocity remains messy (see Keener & Sneyd, referenced at the chapter's end) and not especially illuminating. We omit the complete analytical result here and determine only the maximum reaction velocity.

The maximum reaction velocity V_m' for the allosteric-inhibited reaction is defined as the time-derivative of the product concentration when the reaction is saturated With substrate; that is,

$$\begin{aligned} V_m' &= \lim_{S\to\infty} dP/dt \\ &= k_2 \lim_{S\to\infty} C_1 \end{aligned}$$

With substrate saturation, every enzyme will have its substrate binding site occupied. Enzymes are either bound with only substrate in the complex C_1, or bound together with substrate and inhibitor in the complex C_3. Accordingly, the schematic of the chemical reaction with substrate saturation simplifies to



The equations for C_1 and C_3 with substrate saturation are thus given by

$$\frac{dC_1}{dt} = k_{-3} C_3 - k_3 IC_1 \tag{6.4.4}$$

$$\frac{dC_3}{dt} = k_3 IC_1 - k_{-3} C_3 \tag{6.4.5}$$

and the quasi-equilibrium approximation yields the single independent equation

$$\begin{aligned} C_3 &= (k_3/k_{-3})\, IC_1 \\ &= (I/K_i)\, C_1 \end{aligned} \tag{6.4.6}$$

with $K_i = k_{-3}/k_3$ as before. The equation expressing the conservation of enzyme is given by $E_0 = C_1 + C_3$. This conservation law, together with (6.4.7), permits us to solve for C_1 :

$$C_1 = \frac{E_0}{}$$

$$\cdots = \frac{1}{1 + I/K_i}$$

Therefore, the maximum reaction velocity for the allosteric-inhibited reaction is given by

$$\begin{aligned} V_m' &= \frac{k_2 E_0}{1 + I/K_i} \\ &= \frac{V_m}{1 + I/K_i} \end{aligned}$$

where V_m is the maximum reaction velocity of both the uninhibited and the competitive inhibited reaction. The allosteric inhibitor is thus seen to reduce the maximum velocity of the uninhibited reaction by the factor $(1 + I/K_i)$, which may be large if the concentration of allosteric inhibitor is substantial.

This page titled 6.4: Allosteric Inhibition is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

6.5: Cooperativity

Enzymes and other protein complexes may have multiple binding sites, and when a substrate binds to one of these sites, the other sites may become more active. A well-studied example is the binding of the oxygen molecule to the hemlobin protein. Hemoglobin can bind four molecules of O_2, and when three molecules are bound, the fourth molecule has an increased affinity for binding. We call this cooperativity.

We will model cooperativity by assuming that an enzyme has two separated but indistinguishable binding sites for a substrate S. For example, the enzyme may

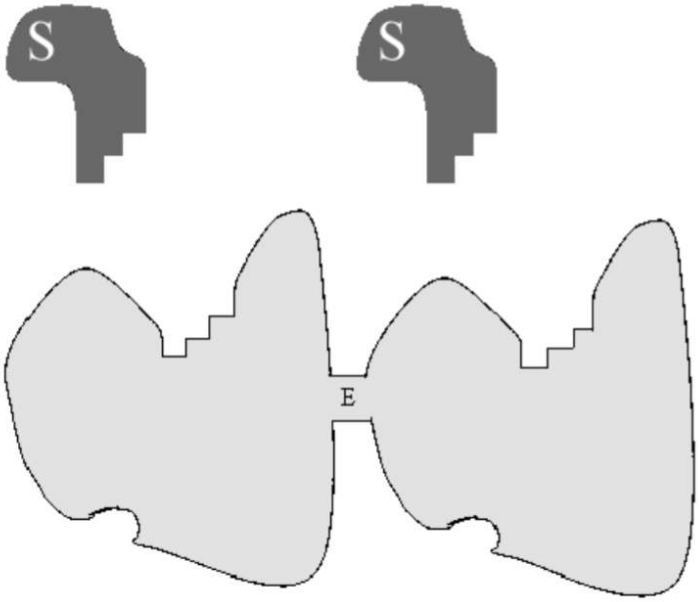
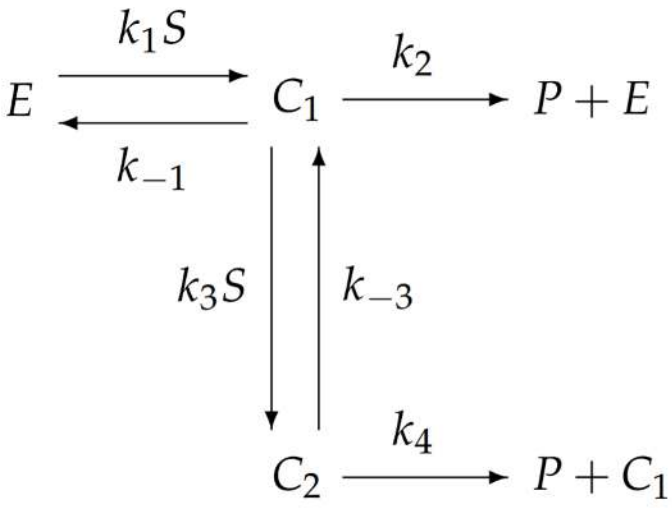


Figure 6.4: Cooperativity.

be a protein dimer, composed of two identical sub-proteins with identical binding sites for S. A cartoon of this enzyme is shown in Fig. 6.4. Because the two binding sites are indistinguishable, we need consider only two complexes: C_1 and C_2, with enzyme bound to one or two substrate molecules, respectively. When the enzyme exhibits cooperativity, the binding of the second substrate molecule has a greater rate constant than the binding of the first. We therefore consider the following reaction:



where cooperativity supposes that $k_1 \ll k_3$. Application of the law of mass action results in

$$\frac{dC_1}{dt} = k_1 SE + (k_{-3} + k_4)\, C_2 - (k_{-1} + k_2 + k_3 S)\, C_1$$

$$\frac{dC_2}{dt} = k_3 S C_1 - (k_{-3} + k_4) C_2$$

Applying the quasi-equilibrium approximation $\dot{C}_1 = \dot{C}_2 = 0$ and the conservation law $E_0 = E + C_1 + C_2$ results in the following system of two equations and two unknowns:

$$(k_{-1} + k_2 + (k_1 + k_3) S) C_1 - (k_{-3} + k_4 - k_1 S) C_2 = k_1 E_0 S, \tag{6.5.1}$$

$$k_3 S C_1 - (k_{-3} + k_4) C_2 = 0. \tag{6.5.2}$$

We divide $(6.5.1)$ by k_1 and $(6.5.2)$ by k_3 and define

$$K_1 = \frac{k_{-1} + k_2}{k_1}, \quad K_2 = \frac{k_{-3} + k_4}{k_3}, \quad \epsilon = k_1 / k_3$$

to obtain

$$(\epsilon K_1 + (1 + \epsilon) S) C_1 - (K_2 - \epsilon S) C_2 = \epsilon E_0 S, \tag{6.5.3}$$

$$S C_1 - K_2 C_2 = 0. \tag{6.5.4}$$

We can subtract (6.5.5) from (6.5.4) and cancel ϵ to obtain

$$(K_1 + S) C_1 + S C_2 = E_0 S.$$

Equations (6.5.5) and (6.5.6) can be solved for C_1 and C_2 :

$$C_1 = \frac{K_2 E_0 S}{K_1 K_2 + K_2 S + S^2} \tag{6.5.5}$$

$$C_2 = \frac{E_0 S^2}{K_1 K_2 + K_2 S + S^2} \tag{6.5.6}$$

so that the reaction velocity is given by

$$\frac{dP}{dt} = k_2 C_1 + k_4 C_2 \tag{6.5.7}$$

$$= \frac{(k_2 K_2 + k_4 S) E_0 S}{K_1 K_2 + K_2 S + S^2} \tag{6.5.8}$$

To illuminate this result, we consider two limiting cases: (i) no cooperativity, where the active sites act independently so that each protein dimer, say, can be considered as two independent protein monomers; (ii) strong cooperativity, where the binding of the second substrate has a much greater rate constant than the binding of the first.

Independent active sites

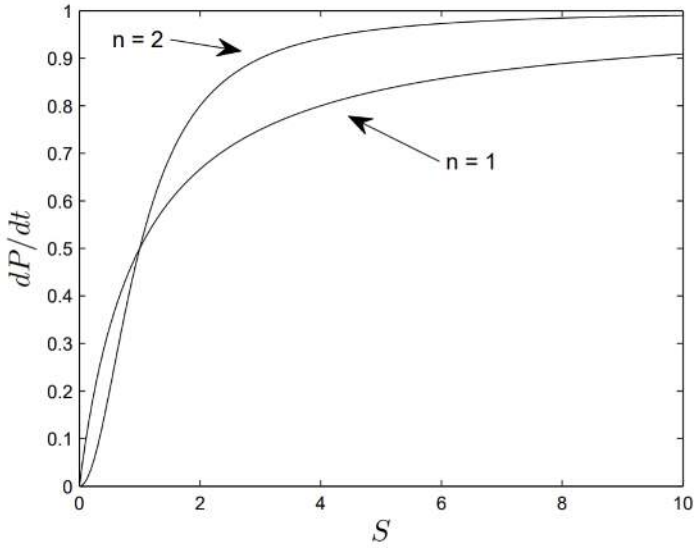


Figure 6.5: The reaction velocity dP/dt as a function of the substrate S. Shown are the solutions to the Hill equation with $V_m = 1, K_m = 1$, and for $n = 1, 2$.

Therefore, for independent active sites, the reaction velocity becomes

$$\frac{dP}{dt} = \frac{(2k_2K_m + 2k_2S)\,E_0S}{K_m^2 + 2K_mS + S^2}$$
$$= \frac{2k_2E_0S}{K_m + S}$$

The reaction velocity for a dimer protein enzyme composed of independent identical monomers is simply double that of a monomer protein enzyme, an intuitively obvious result.

Strong cooperativity

We now assume that after the first substrate binds to the enzyme, the second substrate binds much more easily, so that $k_1 \ll k_3$. The number of enzymes bound to a single substrate molecule should consequently be much less than the number bound to two substrate molecules, resulting in $C_1 \ll C_2$. Dividing (6.5.7) by (6.5.8), this inequality becomes

$$\frac{C_1}{C_2} = \frac{K_2}{S} \ll 1.$$

Dividing the numerator and denominator of $(6.5.9)$ by S^2, we have

$$\frac{dP}{dt} = \frac{(k_2K_2/S + k_4)\,E_0}{(K_1/S)\,(K_2/S) + (K_2/S) + 1}.$$

To take the limit of this expression as $K_2/S \to 0$, we set $K_2/S = 0$ everywhere except in the first term in the denominator, since K_1/S is inversely proportional to k_1 and may go to infinity in this limit. Taking the limit and multiplying the numerator and denominator by S^2,

$$\frac{dP}{dt} = \frac{k_4E_0S^2}{K_1K_2 + S^2}$$

Here, the maximum reaction velocity is $V_m = k_4E_0$, and the modified MichaelisMenten constant is $K_m = \sqrt{K_1K_2}$, so that

$$\frac{dP}{dt} = \frac{V_mS^2}{K_m^2 + S^2}$$

In biochemistry, this reaction velocity is generalized to

$$\frac{dP}{dt} = \frac{V_mS^n}{K_m^n + S^n}$$

known as the Hill equation, and by varying n is used to fit experimental data.

In Fig. 6.5, we have plotted the reaction velocity dP/dt versus S as obtained from the Hill equation with $n = 1$ or 2 . In drawing the figure, we have taken both V_m and K_m equal to unity. It is evident that with increasing n the reaction velocity more rapidly saturates to its maximum value.

This page titled 6.5: Cooperativity is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

7: Sequence Alignment

The software program BLAST (Basic Local Alignment Search Tool) uses sequence alignment algorithms to compare a query sequence against a database to identify other known sequences similar to the query sequence. Often, the annotations attached to the already known sequences yield important biological information about the query sequence. Almost all biologists use BLAST, making sequence alignment one of the most important algorithms of bioinformatics.

The sequence under study can be composed of nucleotides (from the nucleic acids DNA or RNA) or amino acids (from proteins). Nucleic acids chain together four different nucleotides: A,C,T,G for DNA and A, C,U,G for RNA; proteins chain together twenty different amino acids. The sequence of a DNA molecule or of a protein is the linear order of nucleotides or amino acids in a specified direction, defined by the chemistry of the molecule. There is no need for us to know the exact details of the chemistry; it is sufficient to know that a protein has distinguishable ends called the N-terminus and the C-terminus, and that the usual convention is to read the amino acid sequence from the N-terminus to the C-terminus. Specification of the direction is more complicated for a DNA molecule than for a protein molecule because of the double helix structure of DNA, and this will be explained in Section 7.1.

```
CLUSTAL W (1.83) multiple sequence alignment


Human           VHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKV 60
Chimpanzee      VHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKV 60
Rat             VHLTDAEKAAVNGLWGKVNPDDVGGEALGRLLVVYPWTQRYFDSFGDLSSASAIMGNPKV 60
Zebrafish       VEWTDAERTAILGLWGKLNIDEIGPQALSRCLIVYPWTQRYFATFGNLSSPAAIMGNPKV 60
                *. *  *::*: .****:* *::* :**.* *:********:* :**:**:. *:******

Human           KAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGK 120
Chimpanzee      KAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGK 120
Rat             KAHGKKVINAFNDGLKHLDNLKGTFAHLSELHCDKLHVDPENFRLLGNMIVIVLGHHLGK 120
Zebrafish       AAHGRTVMGGLERAIKNMDNVKNTYAALSVMHSEKLHVDPDNFRLLADCITVCAAMKFGQ 120
                 ***:.*:..:. .: ::**:*.*:* ** :*.:******:*****.: :.    . ::*:

Human           E-FTPPVQAAYQKVVAGVANALAHKYH 146
Chimpanzee      E-FTPPVQAAYQKVVAGVANALAHKYH 146
Rat             E-FTPCAQAAFQKVVAGVASALAHKYH 146
Zebrafish       AGFNADVQEAWQKFLAVVVSALCRQYH 147
                  *.. .* *:**.:* *..**.::**
```

Figure 7.1: Multiple alignment of the hemoglobin beta-chain for Human, Chimpanzee, Rat and Zebra fish, obtained using ClustalW.

The basic sequence alignment algorithm aligns two or more sequences to highlight their similarity, inserting a small number of gaps into each sequence (usually denoted by dashes) to align wherever possible identical or similar characters. For instance, Fig 7.1 presents an alignment using the software tool ClustalW of the hemoglobin beta-chain from a human, a chimpanzee, a rat, and a zebrafish. The human and chimpanzee sequences are identical, a consequence of our very close evolutionary relationship. The rat sequence differs from human/chimpanzee at only 27 out of 146 amino acids; we are all mammals. The zebrafish sequence, though clearly related, diverges significantly. Notice the insertion of a gap in each of the mammal sequences at the zebra fish amino acid position 122 . This permits the subsequent zebrafish sequence to better align with the mammal sequences, and implies either an insertion of a new amino acid in fish, or a deletion of an amino acid in mammals. The insertion or deletion of a character in a sequence is called an indel. Mismatches in sequence, such as that occurring between zebrafish and mammals at amino acid positions 2 and 3 is called a mutation. ClustalW places a "* on the last line to denote exact amino acid matches across all sequences, and a′ :′ and ’ ′ ’ to denote chemically similar amino acids across all sequences (each amino acid has characteristic chemical properties, and amino acids can be grouped according to similar properties). In this chapter, we detail the algorithms used to align sequences.

7.1: DNA

7.2: Brute Force Alignment
7.3: Dynamic Programming
7.4: Gaps
7.5: Local Alignments
7.6: Software

This page titled 7: Sequence Alignment is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

7.1: DNA

The minimum you need to know about DNA chem- istry and the genetic code

In one of the most important scientific papers ever published, James Watson and Francis Crick, pictured in Fig. 7.2, determined the structure of DNA using a three- CLUSTAL W (1.83) multiple sequence alignment

Figure 7.1: Multiple alignment of the hemoglobin beta-chain for Human, Chimpanzee, Rat and Zebra fish, obtained using ClustalW.

dimensional molecular model that makes plain the chemical basis of heredity. The DNA molecule consists of two strands wound around each other to form the now famous double helix. Arbitrarily, one strand is labeled by the sequencing group to be the positive strand, and the other the negative strand. The two strands of the DNA molecule bind to each other by base pairing: the bases of one strand pair with the bases of the other strand. Adenine (A) always pairs with thymine (T), and guanine (G) always pairs with cytosine (C) : A with T, G with C. For RNA, T is replaced by uracil (U). When reading the sequence of nucleotides from a single strand, the direction of reading must be specified, and this is possible by referring to the chemical bonds of the DNA backbone. There are of course only two possible directions to read a linear sequence of bases, and these are denoted as $5'-$ to- $3'$ and $3' - to - 5'$. Importantly, the two separate strands of the DNA molecule are oriented in opposite directions. Below is the beginning of the DNA coding sequence for the human hemoglobin beta chain protein discussed earlier:

It is important to realize that there are two unique DNA sequences here, and either one, or even both, can be coding. Reading from $5'-$ to- $3'$, the upper sequence begins 'GTGCACCTG...', while the lower sequence ends'...CAGGTGCAC'. Here, only the upper sequence codes for the human hemoglobin beta chain, and the lower sequence is non-coding.

Figure 7.2: James Watson and Francis Crick posing in front of their DNA model. The original photograph was taken in 1953, the year of discovery, and was recreated in 2003, fifty years later. Francis Crick, the man on the right, died in 2004.

Figure 7.3: The genetic code. codes for a single amino acid. The triplet coding of nucleotides for amino acids is the famous genetic code, shown in Fig. 7.3. Here, the translation to amino acid sequence is 'VHL...', where we have used the genetic code 'GUG' = V, 'CAC'=H, 'CUG' = L. The three out of the twenty amino acids used here are V = Valine, H = Histidine, and L = Leucine

Sequence alignment by brute force

Sequence alignment by dynamic programming

Gap opening and gap extension penalties

7.5 Local alignments

7.6 Software

If you have in hand two or more sequences that you would like to align, there is a choice of software tools available. For relatively short sequences, you can use the LALIGN program for global or local alignments: http://embnet.vital-it.ch/software/LALIGN_form.html

For longer sequences, the BLAST software has a flavor that permits local alignment of two sequences:

http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi

Another useful software for global alignment of two or more long DNA sequences is PipMaker

http://pipmaker.bx.psu.edu/pipmaker/

Multiple global alignments of protein sequences use ClustalW or T-Coffee:

http://www.clustal.org/

Most users of sequence alignment software want to compare a given sequence against a database of sequences. The BLAST software is most widely used, and comes in several versions depending on the type of sequence and database search one is performing:

http://www.ncbi.nlm.nih.gov/BLAST/

This page titled 7.1: DNA is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

7.2: Brute Force Alignment

One (bad) approach to sequence alignment is to align the two sequences in all possible ways, score the alignments with an assumed scoring system, and determine the highest scoring alignment. The problem with this brute-force approach is that the number of possible alignments grows exponentially with sequence length; and for sequences of reasonable length, the computation is already impossible. For example, the number of ways to align two sequences of 50 characters each-a rather small alignment problem-is about 1.5×10^{37}, already an astonishingly large number. It is informative to count the number of possible alignments between two sequences since a similar algorithm is used for sequence alignment.

Suppose we want to align two sequences. Gaps in either sequence are allowed but a gap can not be aligned with a gap. By way of illustration, we demonstrate the three ways that the first character of the upper-case alphabet and the lower-case alphabet may align:

and the five ways in which the first two characters of the upper-case alphabet can align with the first character of the lower-case alphabet:

A recursion relation for the total number of possible alignments of a sequence of i characters with a sequence of j characters may be derived by considering the alignment of the last character. There are three possibilities that we illustrate by assuming the i th character is ' F ' and the j th character is ' d ':

(1) $i-1$ characters of the first sequence are already aligned with $j-1$ characters of the second sequence, and the i th character of the first sequence aligns exactly with the j th character of the second sequence:

(2) $i-1$ characters of the first sequence are aligned with j characters of the second sequence and the i th character of the first sequence aligns with a gap in the second sequence

(3) i characters of the first sequence are aligned with $j-1$ characters of the second sequence and a gap in the first sequence aligns with the j th character of the second sequence iil

$\cdots d$

If $C(i,j)$ is the number of ways to align an i character sequence with a j character sequence, then, from our counting,

$$C(i,j) = C(i-1,j-1) + C(i-1,j) + C(i,j-1)$$

This recursion relation requires boundary conditions. Because there is only one way to align an $i > 0$ character sequence against a zero character sequence (i.e., i characters against i gaps) the boundary conditions are $C(0,j) = C(i,0) = 1$ for all $i, j > 0$ We may also add the additional boundary condition $C(0,0) = 1$, obtained from the known result $C(1,1) = 3$.

Using the recursion relation (7.1), we can construct the following dynamic matrix to count the number of ways to align the two five-character sequences $a_1a_2a_3a_4a_5$ and $b_1b_2b_3b_4b_5$:

	$-$	b_1	b_2	b_3	b_4	b_5
$-$	1	1	1	1	1	1
a_1	1	3	5	7	9	11
a_2	1	5	13	25	41	61
a_3	1	7	25	63	129	231
a_4	1	9	41	129	321	681
a_5	1	11	61	231	681	1683

The size of this dynamic matrix is 6×6, and for convenience we label the rows and columns starting from zero (i.e., row 0 , row $1, \ldots$, row 5). This matrix was constructed by first writing $-a_1a_2a_3a_4a_5$ to the left of the matrix and $-b_1b_2b_3b_4b_5$ above the matrix, then filling in ones across the zeroth row and down the zeroth column to satisfy the boundary conditions, and finally applying the recursion relation directly by going across the first row from left-to-right, the second row from left-to-right, etc. To demonstrate the

filling in of the matrix, we have across the first row: $1+1+1=3, 1+1+3=5, 1+1+5=7$, etc, and across the second row: $1+3+1=5, 3+5+5=13, 5+7+13=25$, etc. Finally, the last element entered gives the number of ways to align two five character sequences: 1683, already a remarkably large number.

It is possible to solve analytically the recursion relation (7.1) for $C(i,j)$ using generating functions. Although the solution method is interesting-and in fact was shown to me by a student-the final analytical result is messy and we omit it here. In general, computation of $C(i,j)$ is best done numerically by constructing the dynamic matrix.

This page titled 7.2: Brute Force Alignment is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

7.3: Dynamic Programming

Two reasonably sized sequences cannot be aligned by brute force. Luckily, there is another algorithm borrowed from computer science, dynamic programming, that makes use of a dynamic matrix.

What is needed is a scoring system to judge the quality of an alignment. The goal is to find the alignment that has the maximum score. We assume that the alignment of character a_i with character b_j has the score $S(a_i, b_j)$. For example, when aligning two DNA sequences, a match (A-A, C-C, T-T, G-G) may be scored as $+2$, and a mismatch $(\mathrm{A}-\mathrm{C}, \mathrm{A}-\mathrm{T}, \mathrm{A}-\mathrm{G}$, etc. $)$ scored as -1. We also assume that an indel (a nucleotide aligned with a gap) is scored as g, with a typical value for DNA alignment being $g=-2$. In the next section, we develop a better and more widely used model for indel scoring that distinguishes gap openings from gap extensions.

Now, let $T(i, j)$ denote the maximum score for aligning a sequence of length i with a sequence of length j. We can compute $T(i, j)$ provided we know $T(i-1, j-1), T(i-1, j)$ and $T(i, j-1)$. Indeed, our logic is similar to that used when counting the total number of alignments. There are again three ways to compute $T(i, j)$: $(1) i-1$ characters of the first sequence are aligned with $j-1$ characters of the second sequence with maximum score $T(i-1, j-1)$, and the i th character of the first sequence aligns with the j th character of the second sequence with updated maximum score $T(i-1, j-1)+S(a_i, b_j)$; $(2) i-1$ characters of the first sequence are aligned with j characters of the second sequence with maximum score $T(i-1, j)$, and the i th character of the first sequence aligns with a gap in the second sequence with updated maximum score $T(i-1, j)+g$, or; (3) i characters of the first sequence are aligned with $j-1$ characters of the second sequence with maximum score $T(i, j-1)$, and a gap in the first sequence aligns with the j th character of the second sequence with updated maximum score $T(i, j-1)+g$. We then compare these three scores and assign $T(i, j)$ to be the maximum; that is,

$$T(i, j)=\max \begin{cases} T(i-1, j-1)+S(a_i, b_j) \\ T(i-1, j)+g \\ T(i, j-1)+g \end{cases}$$

Boundary conditions give the score of aligning a sequence with a null sequence of gaps, so that

$$T(i, 0)=T(0, i)=i g, \quad i>0$$

with $T(0,0)=0$.

The recursion (7.2), together with the boundary conditions (7.3), can be used to construct a dynamic matrix. The score of the best alignment is then given by the last filled-in element of the matrix, which for aligning a sequence of length n with a sequence of length m is $T(n, m)$. Besides this score, however, we also want to determine the alignment itself. The alignment can be obtained by tracing back the path in the dynamic matrix that was followed to compute each matrix element $T(i, j)$. There could be more than one path, so that the best alignment may be degenerate.

Sequence alignment is always done computationally, and there are excellent software tools freely available on the web (see \$7.6). Just to illustrate the dynamic programming algorithm, we compute by hand the dynamic matrix for aligning two short DNA sequences GGAT and GAATT, scoring a match as $+2$, a mismatch as -1 and an indel as -2 :

	–	G	A	A	T	T
–	0	−2	−4	−6	−8	−10
G	−2	2	0	−2	−4	−6
G	−4	0	1	−1	−3	−5
A	−6	−2	2	3	1	−1
T	−8	−4	0	1	5	3

In our hand calculation, the two sequences to be aligned go to the left and above the dynamic matrix, leading with a gap character ' ' '. Row 0 and column 0 are then filled in with the boundary conditions, starting with 0 in position $(0,0)$ and incrementing by the gap penalty -2 across row 0 and down column 0 . The recursion relation (7.2) is then used to fill in the dynamic matrix one row at a time moving from leftto-right and top-to-bottom. To determine the (i, j) matrix element, three numbers must be compared and the maximum taken: (1) inspect the nucleotides to the left of row i and above column j and add $+2$ for a match or -1 for a mismatch to the $(i-1, j-1)$ matrix element; (2) add -2 to the $(i-1, j)$ matrix element; (3) add -2 to the $(i, j-1)$ matrix element. For example, the first computed matrix element 2 at position $(1,1)$ was determined by taking the maximum of $(1) 0+2=2$, since $G-G$ is a

match; $(2) - 2 - 2 = -4; (3) - 2 - 2 = -4$. You can test your understanding of dynamic programming by computing the other matrix elements.

After the matrix is constructed, the traceback algorithm that finds the best alignment starts at the bottom-right element of the matrix, here the $(4, 5)$ matrix element with entry 3. The matrix element used to compute 3 was either at $(4, 4)$ (horizontal move) or at $(3, 4)$ (diagonal move). Having two possibilities implies that the best alignment is degenerate. For now, we arbitrarily choose the diagonal move. We build the alignment from end to beginning with GGAT on top and GAATT on bottom:

T

T

We illustrate our current position in the dynamic matrix by eliminating all the elements that are not on the traceback path and are no longer accessible:

	$-$	G	A	A	T	T
$-$	0	-2	-4	-6	-8	
G	-2	2	0	-2	-4	
G	-4	0	1	-1	-3	
A	-6	-2	2	3	1	
T						3

We start again from the 1 entry at $(3, 4)$. This value came from the 3 entry at $(3, 3)$ by a horizontal move. Therefore, the alignment is extended to

-T

||

TT

where a gap is inserted in the top sequence for a horizontal move. (A gap is inserted in the bottom sequence for a vertical move.) The dynamic matrix now looks like

	$-$	G	A	A	T	T
$-$	0	-2	-4	-6		
G	-2	2	0	-2		
G	-4	0	1	-1		
A	-6	-2	2	3	1	
T						3

Starting again from the 3 entry at $(3, 3)$, this value came from the 1 entry at $(2, 2)$ in a diagonal move, extending the alignment to

A-T

|||

ATT The dynamic matrix now looks like

Continuing in this fashion (try to do this), the final alignment is

where it is customary to represent a matching character with a colon ’ : ’. The traceback path in the dynamic matrix is

If the other degenerate path was initially taken, the final alignment would be

GAATT

and the traceback path would be

The score of both alignments is easily recalculated to be the same, with $2-1+2-2+2=3$ and $2-1+2+2-2=3$

The algorithm for aligning two proteins is similar, except match and mismatch scores depend on the pair of aligning amino acids. With twenty different amino acids found in proteins, the score is represented by a 20×20 substitution matrix. The most commonly used matrices are the PAM series and BLOSUM series of matrices, with BLOSUM62 the commonly used default matrix.

This page titled 7.3: Dynamic Programming is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

7.4: Gaps

Empirical evidence suggests that gaps cluster, in both nucleotide and protein sequences. Clustering is usually modeled by different penalties for gap opening (g_o) and gap extension (g_e), with $g_o < g_e < 0$. For example, the default scoring scheme for the widely used BLASTN software is $+1$ for a nucleotide match, -3 for a nucleotide mismatch, -5 for a gap opening, and -2 for a gap extension.

Having two types of gaps (opening and extension) complicates the dynamic programming algorithm. When an indel is added to an existing alignment, the scoring increment depends on whether the indel is a gap opening or a gap extension. For example, the extended alignment

```
AB   AB−
11   to   111
ab        abc
```

adds a gap opening penalty g_o to the score, whereas

```
A-    A-
II to |||
ab       abc
```

adds a gap extension penalty g_e to the score. The score increment depends not only on the current aligning pair, but also on the previously aligned pair.

The final aligning pair of a sequence of length i with a sequence of length j can be one of three possibilities (top:bottom): (1) $a_i : b_j$; (2) $a_i : -$; (3) $- : b_j$. Only for (1) is the score increment $S(a_i, b_j)$ unambiguous. For (2) or (3), the score increment depends on the presence or absence of indels in the previously aligned characters. For instance, for alignments ending with $a_i : -$, the previously aligned character pair could be one of (i) $a_{i-1} : b_j$, (ii) $- : b_j$, (iii) $a_{i-1} : -$. If the previous aligned character pair was (i) or (ii), the score increment would be the gap opening penalty g_o; if it was (iii), the score increment would be the gap extension penalty g_e.

To remove the ambiguity that occurs with a single dynamic matrix, we need to compute three dynamic matrices simultaneously, with matrix elements denoted by $T(i,j), T_-(i,j)$ and $T^-(i,j)$, corresponding to the three types of aligning pairs. The recursion relations are

$$T(i,j) = \max \begin{cases} T(i-1,j-1) + S(a_i, b_j) \\ T_-(i-1,j-1) + S(a_i, b_j) \\ T^-(i-1,j-1) + S(a_i, b_j) \end{cases}$$

(2) $a_i : -$

$$T_-(i,j) = \max \begin{cases} T(i-1,j) + g_o \\ T_-(i-1,j) + g_e \\ T^-(i-1,j) + g_o \end{cases}$$

(3) $- : b_j$

$$T^-(i,j) = \max \begin{cases} T(i,j-1) + g_o \\ T_-(i,j-1) + g_o \\ T^-(i,j-1) + g_e \end{cases}$$

To align a sequence of length n with a sequence of length m, the best alignment score is the maximum of the scores obtained from the three dynamic matrices:

$$T_{\text{opt}}(n,m) = \max \begin{cases} T(n,m), \\ T_-(n,m), \\ T^-(n,m). \end{cases}$$

The traceback algorithm to find the best alignment proceeds as before by starting with the matrix element corresponding to the best alignment score, $T_{\text{opt}}(n, m)$, and tracing back to the matrix element that determined this score. The optimum alignment is then built up from last-to-first as before, but now switching may occur between the three dynamic matrices.

This page titled 7.4: Gaps is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

7.5: Local Alignments

We have so far discussed how to align two sequences over their entire length, called a global alignment. Often, however, it is more useful to align two sequences over only part of their lengths, called a local alignment. In bioinformatics, the algorithm for global alignment is called "Needleman-Wunsch," and that for local alignment "Smith-Waterman." Local alignments are useful, for instance, when searching a long genome sequence for alignments to a short DNA segment. They are also useful when aligning two protein sequences since proteins can consist of multiple domains, and only a single domain may align.

If for simplicity we consider a constant gap penalty g, then a local alignment can be obtained using the rule

$$T(i,j) = \max \begin{cases} 0, \\ T(i-1,j-1) + S\left(a_i, b_j\right), \\ T(i-1,j) + g, \\ T(i,j-1) + g. \end{cases}$$

After the dynamic matrix is computed using (7.8), the traceback algorithm starts at the matrix element with the highest score, and stops at the first encountered zero score.

If we apply the Smith-Waterman algorithm to locally align the two sequences GGAT and GAATT considered previously, with a match scored as $+2$, a mismatch as -1 and an indel as -2, the dynamic matrix is

	–	G	A	A	T	T
–	0	0	0	0	0	0
G	0	2	0	0	0	0
G	0	2	1	0	0	0
A	0	0	4	3	1	0
T	0	0	2	3	5	3

The traceback algorithm starts at the highest score, here the 5 in matrix element $(4,4)$, and ends at the 0 in matrix element $(0,0)$. The resulting local alignment is

::

GAAT

which has a score of five, larger than the previous global alignment score of three.

This page titled 7.5: Local Alignments is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jeffrey R. Chasnov via source content that was edited to the style and standards of the LibreTexts platform.

Welcome to the Mathematics Library. This Living Library is a principal hub of the LibreTexts project, which is a multi-institutional collaborative venture to develop the next generation of open-access texts to improve postsecondary education at all levels of higher learning. The LibreTexts approach is highly collaborative where an Open Access textbook environment is under constant revision by students, faculty, and outside experts to supplant conventional paper-based books.

Index

A

allosteric inhibition
- 6.3: Competitive Inhibition
- 6.4: Allosteric Inhibition

B

BLAST
- 7: Sequence Alignment

C

competitive inhibition
- 6.3: Competitive Inhibition

complementary cumulative distribution function
- 3.3: Simulation of Population Growth

cooperativity
- 6.5: Cooperativity

D

DNA
- 7.1: DNA

dynamic programming
- 7.3: Dynamic Programming

G

gametic phase disequilibrium
- 5.4: Linkage Equilibrium

genetic code
- 7.1: DNA

genetic drift
- 5.5: Random Genetic Drift

golden ratio
- 2.2: The Golden Ratio Φ

H

haploid genetics
- 5.1: Haploid Genetics

heterosis
- 5.2: Diploid Genetics

Hill equation
- 6.5: Cooperativity

I

indel
- 7.3: Dynamic Programming

interevent time
- 3.3: Simulation of Population Growth

K

Kolmogorov differential equation
- 3.1: A Stochastic Model of Population Growth

L

law of mass action
- 6.1: The Law of Mass Action

linear stability analysis
- 1.2: The Logistic Equation

linkage disequilibrium
- 5.4: Linkage Equilibrium

linkage equilibrium
- 5.4: Linkage Equilibrium

local alignment
- 7.5: Local Alignments

logistic equation
- 1.2: The Logistic Equation

M

Malthusian model
- 1.1: The Malthusian Growth Model

P

polymorphism
- 5.3: Frequency-Dependent Selection

population growth
- 3.3: Simulation of Population Growth

S

sequence alignment
- 7.2: Brute Force Alignment

SIR endemic disease model
- 4.5: The SIR Endemic Disease Model

SIR model
- 4.3: The SIR Epidemic Disease Model

SIS model
- 4.2: The SIS Model

V

virulence
- 4.6: Evolution of Virulence

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Mathematical Biology (Chasnov)

Webpages: 54

All licenses found:

- CC BY 3.0: 94.4% (51 pages)
- Undeclared: 5.6% (3 pages)

By Page

- Mathematical Biology (Chasnov) - *CC BY 3.0*
  - Front Matter - *CC BY 3.0*
    - TitlePage - *CC BY 3.0*
    - InfoPage - *CC BY 3.0*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
    - Preface - *CC BY 3.0*
  - 1: Population Dynamics - *CC BY 3.0*
    - 1.1: The Malthusian Growth Model - *CC BY 3.0*
    - 1.2: The Logistic Equation - *CC BY 3.0*
    - 1.3: A Model of Species Competition - *CC BY 3.0*
    - 1.4: The Lotka-Volterra Predator-Prey Model - *CC BY 3.0*
  - 2: Age-structured Populations - *CC BY 3.0*
    - 2.1: Fibonacci's Rabbits - *CC BY 3.0*
    - 2.2: The Golden Ratio Φ - *CC BY 3.0*
    - 2.3: The Fibonacci Numbers in a Sunflower - *CC BY 3.0*
    - 2.4: Rabbits are an Age-Structured Population - *CC BY 3.0*
    - 2.5: Discrete Age-Structured Populations - *CC BY 3.0*
    - 2.6: Continuous Age-Structured Populations - *CC BY 3.0*
    - 2.7: The Brood Size of a Hermaphroditic Worm - *CC BY 3.0*
  - 3: Stochastic Population Growth - *CC BY 3.0*
    - 3.1: A Stochastic Model of Population Growth - *CC BY 3.0*
    - 3.2: Asymptotics of Large Initial Populations - *CC BY 3.0*
    - 3.3: Simulation of Population Growth - *CC BY 3.0*
  - 4: Infectious Disease Modeling - *CC BY 3.0*
    - 4.1: The SI Model - *CC BY 3.0*
    - 4.2: The SIS Model - *CC BY 3.0*
    - 4.3: The SIR Epidemic Disease Model - *CC BY 3.0*
    - 4.4: Vaccination - *CC BY 3.0*
    - 4.5: The SIR Endemic Disease Model - *CC BY 3.0*
    - 4.6: Evolution of Virulence - *CC BY 3.0*
  - 5: Population Genetics - *CC BY 3.0*
    - 5.1: Haploid Genetics - *CC BY 3.0*
    - 5.2: Diploid Genetics - *CC BY 3.0*
    - 5.3: Frequency-Dependent Selection - *CC BY 3.0*
    - 5.4: Linkage Equilibrium - *CC BY 3.0*
    - 5.5: Random Genetic Drift - *CC BY 3.0*
  - 6: Biochemical Reactions - *CC BY 3.0*
    - 6.1: The Law of Mass Action - *CC BY 3.0*
    - 6.2: Enzyme Kinetics - *CC BY 3.0*
    - 6.3: Competitive Inhibition - *CC BY 3.0*
    - 6.4: Allosteric Inhibition - *CC BY 3.0*
    - 6.5: Cooperativity - *CC BY 3.0*
  - 7: Sequence Alignment - *CC BY 3.0*
    - 7.1: DNA - *CC BY 3.0*
    - 7.2: Brute Force Alignment - *CC BY 3.0*
    - 7.3: Dynamic Programming - *CC BY 3.0*
    - 7.4: Gaps - *CC BY 3.0*
    - 7.5: Local Alignments - *CC BY 3.0*
    - 7.6: Software - *CC BY 3.0*
  - Back Matter - *CC BY 3.0*
    - Index - *CC BY 3.0*
    - Glossary - *CC BY 3.0*
    - Detailed Licensing - *Undeclared*